

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:14:49 ; Search time 76 Seconds
(without alignments)
38.573 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FIHHIFRGIVHAGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	118	100.0	22	23	AAE20172	Fish mast cell ant
2	118	100.0	22	23	ABB07720	Antimicrobial pept
3	87	73.7	22	23	AAE20173	Fish mast cell ant
4	87	73.7	22	23	ABB07721	Antimicrobial pept
5	84	71.2	22	23	AAE20174	Fish mast cell ant
6	84	71.2	22	23	ABB07722	Antimicrobial pept
7	59	50.0	24	23	AAE20177	Antimicrobial pept
8	49	41.5	309	22	ABG25467	Novel human diagno
9	49	41.5	454	22	ABG24377	Novel human diagno
10	48	40.7	509	22	AAU38243	Salmonella typhi c

11	48	40.7	510	19	AAW69358	Pyridine nucleotid
12	48	40.7	510	22	AAU34573	E. coli cellular p
13	48	40.7	883	23	ABB54148	Lactococcus lactis
14	47	39.8	512	21	AAV74582	Neisseria gonorrhe
15	47	39.8	513	21	AAV74583	Neisseria meningit
16	47	39.8	1721	19	AAW52847	A. mediterranei ri
17	46	39.0	480	22	ABB66251	Drosophila melanog
18	46	39.0	480	22	ABB70230	Arabidopsis thalia
19	46	39.0	496	21	AAG27883	Arabidopsis thalia
20	46	39.0	522	21	AAG27882	Arabidopsis thalia
21	46	39.0	687	21	AAG27881	A. thaliana enviro
22	45	38.1	393	21	AAV77947	Micromonospora eve
23	45	38.1	696	22	AAU04874	A. mediterranei ri
24	45	38.1	3413	19	AAW52849	A. mediterranei ri
25	45	38.1	5069	19	AAW52846	Human FCTR3b polyp
26	44.5	37.7	2724	22	AAU08681	Human FCTR3b polyp
27	44.5	37.7	2733	22	AAU08680	S. epidermidis ope
28	44	37.3	271	22	AAG82031	Staphylococcus epi
29	44	37.3	282	23	ABP38310	A delta-12 desatur
30	44	37.3	399	19	AAW84138	Mortierella alpina
31	44	37.3	399	20	AAW95505	M. alpina delta-12
32	44	37.3	399	21	AAV92600	Fungal delta12-des
33	44	37.3	399	21	AAV56046	Amino acid sequenc
34	44	37.3	403	22	AAAB31685	Pouc amino acid se
35	44	37.3	491	23	ABB08382	Novel human diagno
36	44	37.3	765	22	ABG28515	Drosophila melanog
37	43.5	36.9	543	22	ABB59714	Drosophila melanog
38	43.5	36.9	852	22	ABB64019	Protein #2588 enco
39	43	36.4	43	22	ABB20589	Human bone marrow
40	43	36.4	43	22	AAW68359	Peptide #2704 enco
41	43	36.4	43	22	AAW28667	Peptide #2588 enco
42	43	36.4	43	22	AAW03906	Human peptide enco
43	43	36.4	43	23	ABG37920	Fragment of human
44	43	36.4	73	20	AAV02932	Human polypeptide
45	43	36.4	107	22	AAO01861	

ALIGNMENTS

RESULT 1	
AAE20172	standard; peptide: 22 AA.
AAE20172;	
AC	AAE20172;
XX	
DT	18-JUN-2002 (first entry)
XX	
DE	Fish mast cell antimicrobial peptide, piscidin 3 (P3).
XX	
KW	Antimicrobial peptide; endobiotic peptide; mast cell; piscidin; fish;
KW	methicillin; vancomycin; streptogramin; microbial infection; stress;
KW	bacterial disease; fish health; screening; gene therapy; antimicrobial;
KW	tranquilliser; fungicide; protozoacide; antibacterial; antiparasitic.
XX	
OS	Morone chrysops.
OS	Morone saxatilis.
XX	
PN	WO200214345-A2.
XX	
PD	21-FEB-2002.
XX	
PF	13-AUG-2001; 2001WO-US41696.
XX	
PR	15-AUG-2000; 2000US-225354P.
XX	
PA	(UYNC-) UNIV NORTH CAROLINA STATE.
XX	
PI	Noga EJ, Silphaduang U;
XX	
DR	WPI; 2002-269176/31.
XX	
PT	Novel antimicrobial peptides (endobiotic peptides) isolated from mast

PT cells, specifically fish mast cells, useful for treating microbial
PT infections in humans or animals and for reducing antibiotic resistance
PT in bacteria -
XX
PS Claim 4; Page 17; 31pp; English.
XX
CC The invention relates to antimicrobial peptide (endobiotic peptide)
CC isolated from a mast cell. Such peptides obtained from fish mast cells
CC are referred as piscidins. The antimicrobial peptide is useful for
CC treating microbial infection in a subject, and for reducing antibiotic
CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
CC such as Pseudomonas aeruginosa, Escherichia coli, Staphylococcus aureus,
CC Klebsiella pneumoniae, Streptococcus faecalis or Shigella flexneri. These
CC peptides are useful for treating microbial infections in human or animals
CC and for preparing a medicament for treating microbial infections. These
CC peptides are also useful for treating infections caused by microorganisms
CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
CC peptides are useful for treating stress or bacterial disease in fish and
CC monitoring fish health. The inverse relationship between endobiotic
CC levels and stress also allows assessment of freshness of fish food
CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
CC The present sequence is piscidin 3 (P3) peptide isolated from the gills
CC of hybrid striped bass (Morone saxatilis x Morone chrysops).
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 118; DB 23; Length 22;
Best Local Similarity 100.0%; Pred. NO. 4.8e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FIIHIFRGIVHAGRSIGRELTG 22
| | | | | | | | | | | | | | | | | | | | | |
Db 1 FIIHIFRGIVHAGRSIGRELTG 22

RESULT 2
ABB07720
ID ABB07720 standard; peptide; 22 AA.
XX
AC ABB07720;
XX
DT 10-JUN-2002 (first entry)
XX
DE Antimicrobial peptide #1 from hybrid striped bass.
XX
KW Fish; hybrid striped bass; endobiotic; antibiotic; tranquilizer;
KW antimicrobial; antibacterial; fish food product.
XX
OS Morone saxatilis x Morone chrysops.
OS
PN WO200214346-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US41697.
XX
PR 15-AUG-2000; 2000US-225354P.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Noga EJ, Silphaduang U;
XX
DR WPI; 2002-269177/31.
XX
PT Novel antimicrobial compound (endobiotic peptide isolated from fish),
PT useful for monitoring fish health and monitoring freshness of fish, food
PT product and for screening compounds which treat stress in fish -
XX
PS Claim 1; Page 16; 25pp; English.
XX
CC The invention provides antimicrobial peptides (ABB07720-23) isolated from

CC hybrid striped bass. The peptides were purified from the endobiotic
CC family. The endogenous endobiotic peptides exhibiting antimicrobial
CC activity) are useful for monitoring fish health, monitoring the freshness
CC of the fish product such as fresh, refrigerated, or frozen, fish food
CC product, and for screening compounds useful for treating stress in fish.
CC Nucleic acids encoding the antimicrobial peptides are useful for treating
CC stress in a fish. Sequences ABB07720-23 represent specific examples of
CC the antimicrobial peptides.
XX

SQ Sequence 22 AA;

Query Match 100.0%; Score 118; DB 23; Length 22;
Best Local Similarity 100.0%; Pred. NO. 4.8e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FIIHIFRGIVHAGRSIGRELTG 22
| | | | | | | | | | | | | | | | | | | | | |
Db 1 FIIHIFRGIVHAGRSIGRELTG 22

RESULT 3
AAE20173
ID AAE20173 standard; peptide; 22 AA.
XX
AC AAE20173;
XX
DT 18-JUN-2002 (first entry)
XX
DE Fish mast cell antimicrobial peptide, piscidin 1 (P1).
XX

KW Antimicrobial peptide; endobiotic peptide; mast cell; piscidin; fish;
KW methicillin; vancomycin; streptogramin; microbial infection; stress;
KW bacterial disease; fish health; screening; gene therapy; antimicrobial;
KW tranquilizer; fungicide; protozoacide; antibacterial; antiparasitic.
XX

OS Morone chrysops.
OS Morone saxatilis.

PN WO200214345-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US41696.
XX
PR 15-AUG-2000; 2000US-225354P.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Noga EJ, Silphaduang U;
XX
DR WPI; 2002-269176/31.
XX
PT Novel antimicrobial peptides (endobiotic peptides) isolated from mast
PT cells, specifically fish mast cells, useful for treating microbial
PT infections in humans or animals and for reducing antibiotic resistance
PT in bacteria -
XX
PS Claim 4; Page 17; 31pp; English.
XX

CC The invention relates to antimicrobial peptide (endobiotic peptide)
CC isolated from a mast cell. Such peptides obtained from fish mast cells
CC are referred as piscidins. The antimicrobial peptide is useful for
CC treating microbial infection in a subject, and for reducing antibiotic
CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
CC such as Pseudomonas aeruginosa, Escherichia coli, Staphylococcus aureus,
CC Klebsiella pneumoniae, Streptococcus faecalis or Shigella flexneri. These
CC peptides are useful for treating microbial infections in human or animals
CC and for preparing a medicament for treating microbial infections. These
CC peptides are also useful for treating infections caused by microorganisms
CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
CC peptides are useful for treating stress or bacterial disease in fish and
CC monitoring fish health. The inverse relationship between endobiotic
CC levels and stress also allows assessment of freshness of fish food

CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
CC The present sequence is piscidin 1 (P1) peptide isolated from the gills
CC of hybrid striped bass (Morone saxatilis x Morone chrysops).

XX Sequence 22 AA;

Query Match 73.7%; Score 87; DB 23; Length 22;
Best Local Similarity 68.2%; Pred. NO. 4.3e-07;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
1 ||||| 1:1 1:1
Db 1 FFHHIFRGIVHVGKTIHRLVTG 22

RESULT 4

ABB07721 ABB07721 standard; peptide; 22 AA.

XX ABB07721;

DT 10-JUN-2002 (first entry)

DE Antimicrobial peptide #2 from hybrid striped bass.

KW Fish; hybrid striped bass; endobiotic; antibiotic; tranquilizer;

KW antimicrobial; antibacterial; fish food product.

OS Morone saxatilis x Morone chrysops.

PN WO200214346-A2.

PD 21-FEB-2002.

PF 13-AUG-2001; 2001WO-US41697.

PR 15-AUG-2000; 2000US-225354P.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Noga EJ, Silphaduang U;

DR WPI; 2002-269177/31.

PT Novel antimicrobial compound (endobiotic peptide isolated from fish),
PT useful for monitoring fish health and monitoring freshness of fish food

PT product and for screening compounds which treat stress in fish -

PS Claim 1; Page 16; 25pp; English.

CC The invention provides antimicrobial peptides (ABB07720-23) isolated from
CC hybrid striped bass. The peptides were purified from the endobiotic
CC family. The endogenous endobiotic peptides exhibiting antimicrobial
CC activity) are useful for monitoring fish health, monitoring the freshness
CC of the fish product such as fresh, refrigerated, or frozen, fish food
CC product, and for screening compounds useful for treating stress in fish.
CC Nucleic acids encoding the antimicrobial peptides are useful for treating
CC stress in a fish. Sequences ABB07720-23 represent specific examples of
CC the antimicrobial peptides.

SQ Sequence 22 AA;

Query Match 73.7%; Score 87; DB 23; Length 22;
Best Local Similarity 68.2%; Pred. NO. 4.3e-07;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
1 ||||| 1:1 1:1
Db 1 FFHHIFRGIVHVGKTIHRLVTG 22

RESULT 5
AAE20174 AAE20174 standard; peptide; 22 AA.

XX AAE20174;

DT 18-JUN-2002 (first entry)

DE Fish mast cell antimicrobial peptide, piscidin 2 (P2).

KW Antimicrobial peptide; endobiotic peptide; mast cell; piscidin; fish;

KW methicillin; vancomycin; streptogramin; microbial infection; stress;

KW bacterial disease; fish health; screening; gene therapy; antimicrobial;

KW tranquilizer; fungicide; protozoacide; antibacterial; antiparasitic.

OS Morone chrysops.

OS Morone saxatilis.

PN WO200214345-A2.

PD 21-FEB-2002.

PF 13-AUG-2001; 2001WO-US41696.

PR 15-AUG-2000; 2000US-225354P.

PA (UYNC-) UNIV NORTH CAROLINA STATE.

PI Noga EJ, Silphaduang U;

DR WPI; 2002-269176/31.

PT Novel antimicrobial peptides (endobiotic peptides) isolated from mast
PT cells, specifically fish mast cells, useful for treating microbial
PT infections in humans or animals and for reducing antibiotic resistance
PT in bacteria -

PS Claim 4; Page 18; 31pp; English.

CC The invention relates to antimicrobial peptide (endobiotic peptide)
CC isolated from a mast cell. Such peptides obtained from fish mast cells
CC are referred as piscidins. The antimicrobial peptide is useful for
CC treating microbial infection in a subject, and for reducing antibiotic
CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
CC such as Pseudomonas aeruginosa, Escherichia coli, Staphylococcus aureus,
CC Klebsiella pneumoniae, Streptococcus faecalis or Shigella flexneri. These
CC peptides are useful for treating microbial infections in human or animals
CC and for preparing a medicament for treating microbial infections. These
CC peptides are also useful for treating infections caused by microorganisms
CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
CC peptides are useful for treating stress or bacterial disease in fish and
CC monitoring fish health. The inverse relationship between endobiotic
CC levels and stress also allows assessment of freshness of fish food
CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
CC The present sequence is piscidin 2 (P2) peptide isolated from the gills
CC of hybrid striped bass (Morone saxatilis x Morone chrysops).

SQ Sequence 22 AA;

Query Match 71.2%; Score 84; DB 23; Length 22;
Best Local Similarity 63.6%; Pred. NO. 1.3e-06;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
1 ||||| 1:1 1:1
Db 1 FFHHIFRGIVHVGKTIHKLVTG 22

RESULT 6

ABB07722 ABB07722 standard; peptide; 22 AA.

```
XX ABB07722;
AC
XX
XX 10-JUN-2002 (first entry)
DT
XX
XX Antimicrobial peptide #3 from hybrid striped bass.
DE
XX
XX fish; hybrid striped bass; endobiotic; antibiotic; tranquilizer;
KW antimicrobial; antibacterial; fish food product.
XX
OS Morone saxatilis x Morone chrysops.
XX
XX WO200214346-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 13-AUG-2001; 2001WO-US41697.
PF
XX
XX 15-AUG-2000; 2000US-225354P.
PR
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
PA
XX
XX Noga EJ, Silphaduang U;
PI
XX
XX WPI; 2002-269177/31.
DR
XX
XX Novel antimicrobial compound (endobiotic peptide isolated from fish),
PT useful for monitoring fish health and monitoring freshness of fish food
PT product and for screening compounds which treat stress in fish -
PS Claim 1; Page 16; 25pp; English.
XX
XX The invention provides antimicrobial peptides (ABB07720-23) isolated from
CC hybrid striped bass. The peptides were purified from the endobiotic
CC family. The endogenous endobiotic peptides exhibiting antimicrobial
CC activity) are useful for monitoring fish health, monitoring the freshness
CC of the fish product such as fresh, refrigerated, or frozen, fish food
CC product, and for screening compounds useful for treating stress in fish.
CC Nucleic acids encoding the antimicrobial peptides are useful for treating
CC stress in a fish. Sequences ABB07720-23 represent specific examples of
CC the antimicrobial peptides.
XX
SQ Sequence 22 AA;
Query Match 71.2%; Score 84; DB 23; Length 22;
Best Local Similarity 63.6%; Pred. No. 1.3e-06;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 FHHHIFRGIVHAGRSIGRFLTG 22
1 ||||||||| : : ||
Db 1 FHHHIFRGIVHVGKTIHKLVTG 22
RESULT 7
AAE20177 standard; peptide; 24 AA.
XX
XX AAE20177;
AC
XX
XX 18-JUN-2002 (first entry)
DT
XX
XX Antimicrobial peptide analogue.
DE
XX
XX Antimicrobial peptide; endobiotic peptide; mast cell; piscidin;
KW methicillin; vancomycin; streptogramin; microbial infection; stress;
KW bacterial disease; fish health; screening; gene therapy; antimicrobial;
KW tranquilizer; fungicide; protozoacide; antibacterial; antiparasitic.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FH Disulfide-bond 12..13
FT
XX
```

```
PN WO200214345-A2.
XX
XX 21-FEB-2002.
PD
XX
XX 13-AUG-2001; 2001WO-US41696.
PF
XX
XX 15-AUG-2000; 2000US-225354P.
PR
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
PA
XX
XX Noga EJ, Silphaduang U;
PI
XX
XX WPI; 2002-269176/31.
DR
XX
XX Novel antimicrobial peptides (endobiotic peptides) isolated from mast
PT cells, specifically fish mast cells, useful for treating microbial
PT infections in humans or animals and for reducing antibiotic resistance
PT in bacteria -
XX
XX Example 9; Page 24; 31pp; English.
PS
XX
XX The invention relates to antimicrobial peptide (endobiotic peptide)
CC isolated from a mast cell. Such peptides obtained from fish mast cells
CC are referred as piscidins. The antimicrobial peptide is useful for
CC treating microbial infection in a subject, and for reducing antibiotic
CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
CC such as Pseudomonas aeruginosa, Escherichia coli, Staphylococcus aureus,
CC Klebsiella pneumoniae, Streptococcus faecalis or Shigella flexneri. These
CC peptides are useful for treating microbial infections in human or animals
CC and for preparing a medicament for treating microbial infections. These
CC peptides are also useful for treating infections caused by microorganisms
CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
CC peptides are useful for treating stress or bacterial disease in fish and
CC monitoring fish health. The inverse relationship between endobiotic
CC levels and stress also allows assessment of freshness of fish food
CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
XX
XX The present sequence is antimicrobial peptide analogue.
XX
SQ Sequence 24 AA;
Query Match 50.0%; Score 59; DB 23; Length 24;
Best Local Similarity 90.9%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FHHHIFRGIVH 11
1 |||||||||
Db 14 FHHHIFRGIVH 24
RESULT 8
ABG25467 standard; Protein; 309 AA.
XX
XX ABG25467;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #25458.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR
```



```

PR    23-AUG-2000; 2000US-0649167.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Drmanac RT, Liu C, Tang YT;
XX
DR      WPI: 2001-639362/73.
DR      N-PSDB; AAS89654.
XX
PT      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensics, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity -
XX
PS      Claim 20; SEQ ID No 55826; 103pp; English.
XX
CC      The invention relates to isolated polynucleotide (I) and
CC      polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC      polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC      and gene mapping, and in recombinant production of (II). The
CC      polynucleotides are also used in diagnostics as expressed sequence tags
CC      for identifying expressed genes. (I) is useful in gene therapy techniques
CC      to restore normal activity of (II) or to treat disease states involving
CC      (II). (II) is useful for generating antibodies against it, detecting or
CC      quantitating a polypeptide in tissue, as molecular weight markers and as
CC      a food supplement. (II) and its binding partners are useful in medical
CC      imaging of sites expressing (II). (I) and (II) are useful for treating
CC      disorders involving aberrant protein expression or biological activity.
CC      The polypeptide and polynucleotide sequences have applications in
CC      diagnostics, forensics, gene mapping, identification of mutations in
CC      responsible for genetic disorders or other traits to assess biodiversity
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. ABG00010-ABG30377 represent novel human
CC      diagnostic amino acid sequences of the invention.
CC      Note: The sequence data for this patent did not appear in the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence ,   309 AA;

Query Match              41.5%; Score 49; DB 22; Length 309;
Best Local Similarity    33.3%; Pred. No. 9.7;
Matches 11; Conservative 6; Mismatches 4; Indels 12; Gaps 2;

QY      1 FHHIFR-----GIVHA-----GRSIGRELT 21
       1:||||| | | | :::||| 
Db       265 FVHHILRRNGIADGFRLAALTTHGKAVGQYL 297

RESULT 9
ABG24377
ID      ABG24377 standard; Protein; 454 AA.
XX
AC      ABG24377;
DT      18-FEB-2002 (first entry)
XX
DE      Novel human diagnostic protein #24368.
KW      Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder.
OS      Homo sapiens.
XX
PN      WO200175067-A2.
PD      11-OCT-2001.
PF      30-MAR-2001; 2001WO-US08631.
XX
PR      31-MAR-2000; 2000US-0540217.
PR      23-AUG-2000; 2000US-0649167.
XX
```

```

PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS88564.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 54736; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 454 AA;
QY
Query Match 41.5%; Score 49; DB 22; Length 454;
Best Local Similarity 42.3%; Pred. No. 15;
Matches 11; Conservative 2; Mismatches 5; Indels 8; Gaps 1;
QY 3 HHIFRG-----IVHAGRSIGRFL 20
|| || : |||: || ||
Db 426 HHFRRAAQAKRVECVVRAGKEIGHFL 451

RESULT 10
AAU38243
ID AAU38243 standard; Protein; 509 AA.
XX
AC AAU38243;
XX
DT 14-FEB-2002 (first entry)
XX
DE Salmonella typhi cellular proliferation protein #134.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Salmonella typhi.
OS
XX WO200170955-A2.
XX
PN
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.

```

PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS56102.
XX

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Example 3; Seq ID No 13836; 511pp; English.
XX

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pctl_sequences.
XX
SQ Sequence 509 AA;

QY 6 FRGIVHAGRSIGRFLTG 22
: | | | | | | | | | |
DB 139 YRAIVEAAHEFGRFETG 155

Query Match 40.7%; Score 48; DB 22; Length 509;
Best Local Similarity 52.9%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

RESULT 11
AAW69358
ID AAW69358 standard; Protein; 510 AA.
XX
AC AAW69358;
XX
DT 01-DEC-1998 (first entry)
XX
DE Pyridine nucleotide transhydrogenase subunit A.
XX
KW 1,2-propanediol production; intracellular methylglyoxal conversion;
KW unsaturated polyester resin; liquid laundry detergent; antifreeze;
KW de-icing composition; animal feed; pyridine nucleotide transhydrogenase.
XX
OS Unidentified.
XX
PN WO9837204-A1.
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US03271.
XX
PR 19-FEB-1997; 97US-0801344.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX

PI Altaras NE, Cameron DC, Shaw AJ;
XX
XX
DR WPI; 1998-480799/41.
DR N-PSDB; AAV58416.
XX
XX
PT Fermentative production of 1,2-propanediol from simple sugars -
PT using recombinant microorganism that expresses enzymes that convert
PT extracellular methylglyoxal, also new transformants and synthetic
PT operons
XX
PS Disclosure; Page 37-39; 59pp; English.
XX

CC This sequence is the pyridine nucleotide transhydrogenase subunit A. The
CC DNA can be used in a vector used in the method of the invention. The
CC method is for the production of 1,2-propanediol (I), and comprises
CC culturing a recombinant microorganism (A) that expresses enzymes(s) able
CC to convert intracellular methylglyoxal (MG) to (I), in a medium
CC containing a sugar, other than a 6-deoxyhexose, as carbon source. (I) is
CC used in manufacture of unsaturated polyester resins, liquid laundry
CC detergents, pharmaceuticals, cosmetics, antifreeze and de-icing
CC compositions. By-products of the process (carbon dioxide, alcohols and
CC organic acids) may be used e.g. as animal feeds. The method uses
CC inexpensive sugars and provides production of (I) from renewable sources,
CC without: (a) use of propylene oxide as intermediate; (b) generation of
CC toxic waste; and (c) use of high temperature and pressure. Yields of (I)
CC of up to 100 g/l can be achieved.
XX
SQ Sequence 510 AA;

QY 6 FRGIVHAGRSIGRFLTG 22
: | | | | | | | | | |
DB 139 YRAIVEAAHEFGRFETG 155

Query Match 40.7%; Score 48; DB 19; Length 510;
Best Local Similarity 52.9%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

RESULT 12
AAU34573
ID AAU34573 standard; Protein; 510 AA.
XX
AC AAU34573;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #154.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Escherichia coli.
XX
PN WO200170955-A2.
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.

```
DR N-PSDB; AAS52432.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 10166; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 510 AA;
XX
Query Match 40.7%; Score 48; DB 22; Length 510;
Best Local Similarity 52.9%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 6 FRGIVHAGRSIGRFLTG 22
: | | | | | | | | | |
Db 139 YRAIVEAHAHEFGREFTG 155
RESULT 13
ABB54148
ID ABB54148 standard; Protein; 883 AA.
XX
AC ABB54148;
XX
DT 16-MAY-2002 (first entry)
XX
DE Lactococcus lactis protein yieh.
XX
KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS Lactococcus lactis IL1403.
XX
PN FR2807446-A1.
XX
PD 12-OCT-2001.
XX
PF 11-APR-2000; 2000FR-0004630.
XX
PR 11-APR-2000; 2000FR-0004630.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
DR WPI; 2002-043418/06.
XX
PS Claim 6; SEQ ID No 850; 2504pp; French.
XX
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
```

CC	nucleic acid sequence is useful in the detection and/or amplification of
CC	nucleic acid sequence, particularly to identify <i>Lactococcus lactis</i> or
CC	related species. The proteins of the invention are useful for the
CC	biosynthesis or biodegradation of a composition of interest. The
CC	invention helps research in lactic bacteria, particularly useful in the
CC	production of yogurt and cheese.
CC	Note: The sequence data for this patent is based on equivalent patent
CC	WO200177334 (published 18-OCT-2001) which is available in electronic
CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences .
XX	
SD	Sequence 883 AA:
QY	2 IHIFRGIVHAGRSIGRFLT 21
Db	11 : 1 1 : 1 1 1 51 IHSLYRNILHSGSGSGLYLT 70
RESULT 14	
AAV74582	
ID	AAV74582 standard; Protein; 512 AA.
XX	
AC	AAV74582;
XX	
DT	21-MAR-2000 (first entry)
XX	
DE	Neisseria gonorrhoeae ORF 155 protein sequence SEQ ID NO:638.
XX	
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX	antibacterial; gene therapy.
XX	
OS	Neisseria gonorrhoeae.
XX	
PN	WO9957280-A2.
XX	
PD	11-NOV-1999.
XX	
PF	30-APR-1999; 99WO-US09346.
XX	
PR	01-MAY-1998; 98US-0083758.
PR	31-JUL-1998; 98US-0094869.
PR	02-SEP-1998; 98US-0098994.
PR	02-SEP-1998; 98US-0099062.
PR	09-OCT-1998; 98US-0103749.
PR	09-OCT-1998; 98US-0103794.
PR	09-OCT-1998; 98US-0103796.
PR	25-FEB-1999; 99US-0121528.
XX	
PA	(CHTR) CHIRON CORP.
PA	(GENO-) INST GENOMIC RES.
XX	
PI	Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
PI	Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI	Tettelin H, Venter JC;
XX	
DR	WPI, 2000-062150/05.
DR	N-PSDB; AAZ53344.
XX	
PT	Novel Neisserial polypeptides predicted to be useful antigens for
PT	vaccines and diagnostics
XX	
PS	Claim 2; Page 444; 1453pp; English.
XX	
CC	AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAV74253 to AAV75941
CC	represent novel <i>Neisseria meningitidis</i> and <i>N. gonorrhoeae</i> polynucleotides
CC	and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC	PCR primers used in the exemplification of the present invention. The
CC	polypeptides, the polynucleotides, antibodies and compositions of
CC	the invention can be used as vaccines, as diagnostic reagents, and as

CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.

XX SQ Sequence 512 AA;

Query Match 39.8%; Score 47; DB 21; Length 512;

Best Local Similarity 41.2%; Pred. No. 36;

Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 6 FRGIVHAGRSIGRFLTG 22

DB 139 YRAVIEAANAFGRFFTG 155

RESULT 15

AAV74583 standard; Protein; 513 AA.

AC AAV74583;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 155 protein sequence SEQ ID NO:640.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

XX OS Neisseria meningitidis.

PN WO9957280-A2.

PD 11-NOV-1999.

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098994.

PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.

XX (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
PI Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

DR N-PSDB; AAZ53345.

XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics

XX Claim 2; Page 445; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to

CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.

XX SQ Sequence 513 AA;

Query Match 39.8%; Score 47; DB 21; Length 513;

Best Local Similarity 41.2%; Pred. No. 36;

Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 6 FRGIVHAGRSIGRFLTG 22

DB 139 YRAVIEAANAFGRFFTG 155

Search completed: May 17, 2003, 14:25:09
Job time : 78 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:23:54 ; Search time 28 Seconds
(without alignments)
23.118 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FHHHFRGIVHAGRSTGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	40.7	510	3 US-08-801-344-8	Sequence 8, Appli
2	48	40.7	510	4 US-09-498-599-8	Sequence 8, Appli
3	44	37.3	282	4 US-09-134-001C-3155	Sequence 3155, Ap
4	44	37.3	399	2 US-08-834-655-4	Sequence 4, Appli
5	44	37.3	399	3 US-08-834-033A-4	Sequence 4, Appli
6	44	37.3	399	4 US-09-363-574-4	Sequence 4, Appli
7	44	37.3	399	4 US-09-363-526-4	Sequence 4, Appli
8	44	37.3	399	4 US-09-330-235-20	Sequence 20, Appli
9	43	36.4	73	4 US-09-227-357-586	Sequence 586, App
10	43	36.4	471	4 US-09-134-001C-4904	Sequence 4904, Ap
11	42	35.6	362	3 US-08-513-974B-374	Sequence 374, App
12	41	34.7	885	1 US-08-042-747A-8	Sequence 8, Appli
13	41	34.7	885	3 US-08-804-439A-23	Sequence 23, Appli
14	41	34.7	885	3 US-08-720-229-23	Sequence 23, Appli
15	41	34.7	891	1 US-08-042-747A-6	Sequence 6, Appli
16	41	34.7	2296	2 US-08-286-819A-27	Sequence 27, Appli
17	41	34.7	2296	3 US-08-980-357-27	Sequence 27, Appli
18	40.5	34.3	170	3 US-09-130-663-2	Sequence 2, Appli
19	40.5	34.3	170	4 US-09-432-335-2	Sequence 2, Appli
20	40.5	34.3	170	4 US-09-614-022-2	Sequence 2, Appli
21	40	33.9	405	4 US-09-134-001C-4999	Sequence 4999, Ap
22	40	33.9	690	4 US-09-134-001C-4568	Sequence 4568, Ap
23	39.5	33.5	354	4 US-09-134-001C-3374	Sequence 3374, Ap
24	39	33.1	187	4 US-09-443-184-59	Sequence 59, Appli
25	39	33.1	187	2 US-08-713-825-1	Sequence 1, Appli
26	39	33.1	187	3 US-09-199-842-1	Sequence 1, Appli
27	39	33.1	327	1 US-07-867-105B-1	Sequence 1, Appli

28	39	33.1	376	4 US-09-056-556-202	Sequence 202, App
29	39	33.1	376	4 US-09-072-596-197	Sequence 197, App
30	39	33.1	567	2 US-08-504-459-2	Sequence 2, Appli
31	38	32.2	125	1 US-08-249-013-7	Sequence 7, Appli
32	38	32.2	125	2 US-08-886-863-7	Sequence 7, Appli
33	38	32.2	125	4 US-09-175-229-7	Sequence 7, Appli
34	38	32.2	125	5 PCT-US95-06764-7	Sequence 7, Appli
35	38	32.2	159	4 US-09-280-839-3	Sequence 3, Appli
36	38	32.2	166	3 US-09-081-180-4	Sequence 4, Appli
37	38	32.2	166	3 US-09-040-786-4	Sequence 4, Appli
38	38	32.2	436	1 US-08-674-168-6	Sequence 6, Appli
39	38	32.2	622	4 US-09-311-626B-4	Sequence 4, Appli
40	38	32.2	865	1 US-07-803-633A-13	Sequence 13, Appli
41	37	31.4	129	3 US-08-513-974B-327	Sequence 327, App
42	37	31.4	357	4 US-09-149-476-755	Sequence 755, App
43	37	31.4	411	4 US-08-887-534A-80	Sequence 80, Appli
44	37	31.4	423	4 US-09-134-001C-5210	Sequence 5210, Ap
45	37	31.4	436	1 US-08-146-010A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-801-344-8
Sequence 8, Application US/08801344
Patent No. 6087140
GENERAL INFORMATION:
APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.
TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
TITLE OF INVENTION: FROM SUGAR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dewitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,344
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 09820.037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Pyridine nucleotide transhydrogenase, subunit
US-08-801-344-8
Query Match 40.7%; Score 48; DB 3; Length 510;
Best Local Similarity 52.9%; Pred. No. 7.1;

Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIVHAGRSIGRFLTG 22
:| | | | | | | | | |
Db 139 YRAIVEAAHEFGRRFTG 155

RESULT 2

US-09-498-599-8
; Sequence 8, Application US/09498599
; Patent No. 6303352

GENERAL INFORMATION:

APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.
TITLE OF INVENTION: MICROBIAL PRODUCTION OF
TITLE OF INVENTION: 1,2-PROPANEDIOL FROM SUGAR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,599

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sara, Charles S.

REFERENCE/DOCKET NUMBER: 09820.037

TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-831-2100

TELEFAX: 608-831-2106

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 510 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Pyridine nucleotide transhydrogenase, subunit

ORGANISM: A

Query Match

Best Local Similarity 40.7%; Score 48; DB 4; Length 510;

Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIVHAGRSIGRFLTG 22
:| | | | | | | | | |

Db 139 YRAIVEAAHEFGRRFTG 155

RESULT 3

US-09-134-001C-3155

; Sequence 3155, Application US/09134001C

; Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3155
;; LENGTH: 282
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3155

Query Match

Best Local Similarity 37.3%; Score 44; DB 4; Length 282;

Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 FTHHIFRGIVHAGRSIG 17
|::|:| | | | |

Db 188 FLYHVEQYKAGTGLG 204

RESULT 4

US-08-834-655-4

; Sequence 4, Application US/08834655

; Patent No. 5968809

GENERAL INFORMATION:

APPLICANT: KNUTZON, DEBORAH

APPLICANT: MURKERUT, PRADIP

APPLICANT: HUANG, YUNG-SHENG

APPLICANT: THURMOND, JENNIFER

APPLICANT: CHAUDHARY, SUNITA

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS

TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: RAE-VENTER LAW GROUP, P.C.

STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/834,655

FILING DATE: 11-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: RAE-VENTER, BARBARA

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: CGNE.124.00US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 328-4400

TELEFAX: (650) 328-4477

TELEX: N/A

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 399 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-834-655-4

Query Match 37.3%; Score 44; DB 2; Length 399;

Best Local Similarity 63.6%; Pred. No. 23;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FTHHIFRGIVH 11

Db 326 FLDHMFHGIHV 336

RESULT 5

US-08-834-033A-4
; Sequence 4, Application US/08834033A
; Patent No. 6075183
; GENERAL INFORMATION:
; APPLICANT: KNU TZON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,033A
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-300.USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-834-033A-4

Query Match 37.3%; Score 44; DB 3; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FHHHFRGIHV 11
1:1:11111
Db 326 FLDHMFHGIHV 336

RESULT 6

US-09-363-574-4
; Sequence 4, Application US/09363574
; Patent No. 6136574
; GENERAL INFORMATION:
; APPLICANT: KNU TZON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:

ADDRESSEE: LIMBACH AND LIMBACH L.L.P.
STREET: 2001 FERRY BUILDING
CITY: SAN FRANCISCO
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/363,574
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WARD, MICHAEL R.
REGISTRATION NUMBER: 38,651
REFERENCE/DOCKET NUMBER: CGAB-202 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-363-574-4

Query Match 37.3%; Score 44; DB 4; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FHHHFRGIHV 11
1:1:11111
Db 326 FLDHMFHGIHV 336

RESULT 7

US-09-363-526-4
; Sequence 4, Application US/09363526
; Patent No. 6410288
; GENERAL INFORMATION:
; APPLICANT: KNU TZON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,526
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651

REFERENCE/DOCKET NUMBER: CGAB-201 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-363-526-4

Query Match 37.3%; Score 44; DB 4; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FIHHIFRGIVH 11
|:|:|
Db 326 FLDMFHGIVH 336

RESULT 8
US-09-330-235-20
; Sequence 20, Application US/09330235
; Patent No. 6453018
; GENERAL INFORMATION:
; APPLICANT: Knutzon, Debbie
; TITLE OF INVENTION: POLYUNSATURATED FATTY ACIDS IN PLANTS
; FILE REFERENCE: MOCO.156.00US
; CURRENT APPLICATION NUMBER: US/09/330,235
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: 60/089,043
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-330-235-20

Query Match 37.3%; Score 44; DB 4; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FIHHIFRGIVH 11
|:|:|
Db 326 FLDMFHGIVH 336

RESULT 9
US-09-227-357-586
; Sequence 586, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 586
LENGTH: 73
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (66)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-227-357-586

Query Match 36.4%; Score 43; DB 4; Length 73;
Best Local Similarity 43.8%; Pred. No. 5.6;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FIHHIFRGIVHGRSI 16
| | | : | | |

Db 55 FAHLLILVAHXGCSI 70

RESULT 10
US-09-134-001C-4904
; Sequence 4904, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucelte-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4904
; LENGTH: 471
; TYPE: PR1
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4904

Query Match 36.4%; Score 43; DB 4; Length 471;
Best Local Similarity 80.0%; Pred. NO. 40;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 IFRGIVHAGR 14
|:|||||11
Db 361 IYRGIVHAGR 370

RESULT 11
US-08-513-974B-374
; Sequence 374, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Jetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 7-007177
;; FILING DATE: 20-JAN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-326611
;; FILING DATE: 28-DEC-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-270017
;; FILING DATE: 02-NOV-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-236357
;; FILING DATE: 30-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-236356
;; FILING DATE: 30-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-189274
;; FILING DATE: 11-AUG-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-189273
;; FILING DATE: 11-AUG-1945
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-189272
;; FILING DATE: 11-AUG-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Resnick, David S.
;; REGISTRATION NUMBER: 34,235
;; REFERENCE/DOCKET NUMBER: 45753
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-523-3400
;; TELEFAX: 617-523-6440
;; INFORMATION FOR SEQ ID NO: 374:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 362 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-513-974B-374

Query Match 35.6%; Score 42; DB 3; Length 362;
Best Local Similarity 46.7%; Pred. NO. 43;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 HIFRGIVHAGRSIGR 18
|:|:|11:|:|11
Db 137 HRYTGIVHPLKSLGR 151

RESULT 12
US-08-042-747A-8
; Sequence 8, Application US/08042747A
; Patent No. 5487969
; GENERAL INFORMATION:
; APPLICANT: Eberle, Richard
; APPLICANT: Black, Darla
; APPLICANT: Scinicariello, Franco
; APPLICANT: Hilliard, Julia K.
; TITLE OF INVENTION: Cloning and Amplification of Monkey B
; TITLE OF INVENTION: Virus Genes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Cox & Smith Incorporated
STREET: 112 East Pecan street, Suite 2000
CITY: San Antonio
STATE: Texas
COUNTRY: USA
ZIP: 78205
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/042,747A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haymond, W. Bradley
REGISTRATION NUMBER: 35186
REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395
TELEX: 767609
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-042-747A-8

Query Match 34.7%; Score 41; DB 1; Length 885;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 IHHIFRGIHVHAGRSIGRFLTG 22
Db 720 LYSFEEGLGDVGRVAVGMG 740

RESULT 13

US-08-804-439A-23
Sequence 23, Application US/08804439A
Patent No. 6015565

GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kuit
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Ste 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439A
FILING DATE: February 21, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09176/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
TELEX:

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-804-439A-23

Query Match 34.7%; Score 41; DB 3; Length 885;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 IHHIFRGIHVHAGRSIGRFLTG 22
Db 720 LYSFEEGLGDVGRVAVGMG 740

RESULT 14

US-08-720-229-23
Sequence 23, Application US/08720229
Patent No. 6022542

GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kuit
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-23

Query Match 34.7%; Score 41; DB 3; Length 885;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 IHHIFRGIHVHAGRSIGRFLTG 22
Db 720 LYSFEEGLGDVGRVAVGMG 740

RESULT 15

US-08-042-747A-6
Sequence 6, Application US/08042747A
Patent No. 5487969

GENERAL INFORMATION:

APPLICANT: Eberle, Richard
APPLICANT: Black, Darla
APPLICANT: Scinicariello, Franco
APPLICANT: Hilliard, Julia K.
TITLE OF INVENTION: Cloning and Amplification of Monkey B
TITLE OF INVENTION: Virus Genes

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cox & Smith Incorporated
STREET: 112 East Pecan Street, Suite 2000
CITY: San Antonio
STATE: Texas
COUNTRY: USA
ZIP: 78205
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/042,747A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haymond, W. Bradley
REGISTRATION NUMBER: 35186
REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395
TELEX: 767609
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 891 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-042-747A-6
Query Match 34.7%; Score 41; DB 1; Length 891;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 2 IHHIFRGIVHAGRSIGRFLTG 22
DB 722 LYSFEEGLGDYGRAVGKVMG 742

Search completed: May 17, 2003, 14:28:37
Job time : 30 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:27:19 ; Search time 54 Seconds
(without alignments)
39.295 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FIHHIFRGIVHAGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	100.0	22	US-09-929-788-1	Sequence 1, Appli
2	87	73.7	22	US-09-929-788-2	Sequence 2, Appli
3	84	71.2	22	US-09-929-788-3	Sequence 3, Appli
4	80	67.8	23	US-10-076-816-42	Sequence 42, Appli
5	48	40.7	509	US-09-815-242-13836	Sequence 13836, A
6	48	40.7	510	US-09-815-242-10166	Sequence 10166, A
7	44.5	37.7	1737	US-09-808-602-83	Sequence 83, Appli
8	44.5	37.7	1737	US-09-800-198-71	Sequence 71, Appli
9	44.5	37.7	2724	US-09-808-602-13	Sequence 13, Appli
10	44.5	37.7	2724	US-09-800-198-13	Sequence 13, Appli
11	44.5	37.7	2733	US-09-808-602-8	Sequence 8, Appli
12	44.5	37.7	2733	US-09-800-198-8	Sequence 8, Appli
13	44.5	37.7	2802	US-09-808-602-81	Sequence 81, Appli
14	44.5	37.7	2802	US-09-800-198-69	Sequence 69, Appli
15	44	37.3	491	US-09-997-664-3	Sequence 3, Appli
16	43	36.4	43	US-09-864-761-35887	Sequence 35887, A
17	43	36.4	73	US-09-983-802-586	Sequence 586, App
18	43	36.4	1436	US-09-815-242-5566	Sequence 5566, Ap
19	43	36.4	1442	US-09-815-242-12321	Sequence 12321, A

20	42	35.6	362	9	US-10-092-135-3	Sequence 3, Appli
21	42	35.6	362	9	US-10-092-135-4	Sequence 4, Appli
22	42	35.6	362	9	US-09-779-679-28	Sequence 28, Appli
23	41	34.7	169	9	US-09-746-783-80	Sequence 80, Appli
24	41	34.7	247	9	US-10-028-072-514	Sequence 514, App
25	41	34.7	247	9	US-10-174-590-604	Sequence 604, App
26	41	34.7	247	9	US-10-176-758-604	Sequence 604, App
27	41	34.7	247	9	US-10-175-737-604	Sequence 604, App
28	41	34.7	247	9	US-10-121-049-514	Sequence 514, App
29	41	34.7	247	9	US-10-123-904-514	Sequence 514, App
30	41	34.7	247	9	US-10-140-470-514	Sequence 514, App
31	41	34.7	247	9	US-10-173-706-604	Sequence 604, App
32	41	34.7	247	9	US-10-175-738-604	Sequence 604, App
33	41	34.7	247	9	US-10-175-752-604	Sequence 604, App
34	41	34.7	247	9	US-10-176-482-604	Sequence 604, App
35	41	34.7	247	9	US-10-176-757-604	Sequence 604, App
36	41	34.7	247	9	US-10-176-913-604	Sequence 604, App
37	41	34.7	247	9	US-10-180-552-604	Sequence 604, App
38	41	34.7	247	9	US-10-180-557-604	Sequence 604, App
39	41	34.7	247	9	US-10-173-700-604	Sequence 604, App
40	41	34.7	247	9	US-10-174-572-604	Sequence 604, App
41	41	34.7	247	9	US-10-174-579-604	Sequence 604, App
42	41	34.7	247	9	US-10-174-582-604	Sequence 604, App
43	41	34.7	247	9	US-10-174-588-604	Sequence 604, App
44	41	34.7	247	9	US-10-175-739-604	Sequence 604, App
45	41	34.7	247	9	US-10-175-740-604	Sequence 604, App

ALIGNMENTS

RESULT 1
US-09-929-788-1
; Sequence 1, Application US/09929788
; Publication No. US20030083247A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030083247A1a, Edward J.
; APPLICANT: Silphaduang, Umaborn
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES ISOLATED FROM FISH
; FILE REFERENCE: 5051.519
; CURRENT APPLICATION NUMBER: US/09/929, 788
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,354
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Morone saxatilis x Morone chrysops
; US-09-929-788-1

Query Match 100.0%; Score 118; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 8.5e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FIHHIFRGIVHAGRSIGRFLTG 22
Db 1 FIHHIFRGIVHAGRSIGRFLTG 22

RESULT 2
US-09-929-788-2
; Sequence 2, Application US/09929788
; Publication No. US20030083247A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030083247A1a, Edward J.
; APPLICANT: Silphaduang, Umaborn
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES ISOLATED FROM FISH
; FILE REFERENCE: 5051.519
; CURRENT APPLICATION NUMBER: US/09/929, 788
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,354

;; PRIOR FILING DATE: 2000-08-15
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 22
;; TYPE: PRT
;; ORGANISM: Morone saxatilis x Morone chrysops
US-09-929-788-2

Query Match 73.7%; Score 87; DB 9; Length 22;
Best Local Similarity 68.2%; Pred. No. 2.7e-06;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
Db 1 FHHIFRGIVHVGKTIHRLVTG 22

RESULT 3

US-09-929-788-3
; Sequence 3, Application US/09929788
; Publication No. US20030083247A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030083247A1a, Edward J.
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES ISOLATED FROM FISH
; FILE REFERENCE: 5051.519
; CURRENT APPLICATION NUMBER: US/09/929,788
; PRIOR FILING DATE: 2001-08-14
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Morone saxatilis x Morone chrysops
US-09-929-788-3

Query Match 71.2%; Score 84; DB 9; Length 22;
Best Local Similarity 63.6%; Pred. No. 7.2e-06;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
Db 1 FHHIFRGIVHVGKTIHKLVTG 22

RESULT 4

US-10-076-816-42
; Sequence 42, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 23
; TYPE: PRT

;; ORGANISM: Morone sp.
US-10-076-816-42

Query Match 67.8%; Score 80; DB 9; Length 23;
Best Local Similarity 63.6%; Pred. No. 2.9e-05;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
Db 1 FHHIFRGIVHVGKTIHDLVTG 22

RESULT 5

US-09-815-242-13836
; Sequence 13836, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13836
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(509)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13836

Query Match 40.7%; Score 48; DB 10; Length 509;
Best Local Similarity 52.9%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIVHAGRSIGRFLTG 22
Db 139 YRAIVEAHERGFRFTG 155

RESULT 6

US-09-815-242-10166
; Sequence 10166, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.


```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10166
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10166
```

```

Query Match          40.7%; Score 48; DB 10; Length 510;
Best Local Similarity 52.9%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 6 FRGIVHAGRSIGRFLT 22
Db 139 YRAIVEAAHEFGRFETG 155
```

```

RESULT 7
US-09-808-602-83
; Sequence 83, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-83
```

```

Query Match          37.7%; Score 44.5; DB 9; Length 1737;
Best Local Similarity 45.5%; Pred. No. 3.7e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;
```

```

QY 3 HHIFR---GIYHAGRSIGRFLT 21
Db 147 HHILNVKSGILHKGTGENQFLT 168
```

```

RESULT 8
US-09-800-198-71
; Sequence 71, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-198-71
```

```

Query Match          37.7%; Score 44.5; DB 9; Length 1737;
Best Local Similarity 45.5%; Pred. No. 3.7e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;
QY 3 HHIFR---GIYHAGRSIGRFLT 21
Db 147 HHILNVKSGILHKGTGENQFLT 168
```

```

RESULT 9
US-09-808-602-13
; Sequence 13, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-13
```

```

Query Match          37.7%; Score 44.5; DB 9; Length 2724;
Best Local Similarity 45.5%; Pred. No. 5.9e+02;
```

Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHIFR--GIVHAGRSIGRFLT 21
||| ||:| | :|||
Db 1134 HHILNVKSGILHKGTGENQFLT 1155

RESULT 10
US-09-800-198-13
; Sequence 13, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-198-13

Query Match 37.7%; Score 44.5; DB 9; Length 2724;
Best Local Similarity 45.5%; Pred. No. 5.9e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHIFR--GIVHAGRSIGRFLT 21
||| ||:| | :|||
Db 1134 HHILNVKSGILHKGTGENQFLT 1155

RESULT 11
US-09-808-602-8
; Sequence 8, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2733
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-8

Query Match 37.7%; Score 44.5; DB 9; Length 2733;

Best Local Similarity 45.5%; Pred. No. 5.9e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHIFR--GIVHAGRSIGRFLT 21
||| ||:| | :|||
Db 1143 HHILNVKSGILHKGTGENQFLT 1164

RESULT 12
US-09-800-198-8
; Sequence 8, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2733
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-198-8

Query Match 37.7%; Score 44.5; DB 9; Length 2733;
Best Local Similarity 45.5%; Pred. No. 5.9e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHIFR--GIVHAGRSIGRFLT 21
||| ||:| | :|||
Db 1143 HHILNVKSGILHKGTGENQFLT 1164

RESULT 13
US-09-808-602-81
; Sequence 81, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 2802
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-808-602-81

Query Match 37.7%; Score 44.5; DB 9; Length 2802;
Best Local Similarity 40.9%; Pred. No. 6.1e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

OY 3 HHIFR--GIVHAGRSIGRFLT 21
11: 11:11 :111
Db 1212 HHVLNVKSGILHKNGENQFLT 1233

RESULT 14

US-09-800-198-69
; Sequence 69, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkels, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
; LENGTH: 2802
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-800-198-69

Query Match 37.7%; Score 44.5; DB 9; Length 2802;
Best Local Similarity 40.9%; Pred. No. 6.1e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

OY 3 HHIFR--GIVHAGRSIGRFLT 21
11: 11:11 :111
Db 1212 HHVLNVKSGILHKNGENQFLT 1233

RESULT 15

US-09-997-664-3
; Sequence 3, Application US/09997664
; Patent No. US20020151003A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arie
; APPLICANT: Catermole, Monica
; APPLICANT: Gatenby, Anthony A.
; APPLICANT: Gibson, Katherine J.
; APPLICANT: Ramos-Gonzalez, Isabel
; APPLICANT: Ramos, Juan
; APPLICANT: Sariaslani, Sima
; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
; TITLE OF INVENTION: Pseudomonas and Agrobacterium
; FILE REFERENCE: BC1018 US CIP
; CURRENT APPLICATION NUMBER: US/09/997,664
; PRIOR APPLICATION NUMBER: 09/585,174
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Pseudomonas mendocina KR-1
US-09-997-664-3

Query Match 37.3%; Score 44; DB 10; Length 491;

Best Local Similarity 35.7%; Pred. No. 1.2e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

OY 1 FIHH-----IFRGIVHAGRSIGRFLTG 22
11: 11:11 :111
Db 222 FVEHPVPALISFTGSTQVGRNIGRIASG 249

Search completed: May 17, 2003, 14:36:55
Job time : 56 secs

A:Map position: circular chromosome

C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogenase

Query Match

Best Local Similarity 44.9%; Score 53; DB 2; Length 524;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 FRGIVHAGRSIGRFLTG 22
Db 141 YRAVIEAGNNFGRRFTG 157

RESULT 3

VCGBEBH

glycoprotein B precursor - bovine herpesvirus 2 (strain BMV)

C:Species: bovine herpesvirus 2

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

C:Accession: C29242

R:Hamerschmidt, W.; Contraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.

Virolgy 165, 388-405, 1988

A:Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus 2

A:Reference number: A94381; MUID:88306231; PMID:2841793

A:Accession: C29242

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-917 <HAM>

A:Cross-references: GB:M21628; NID:g330752; PIDN:AAA46053.1; PID:g330753

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-917/Product: glycoprotein B #status predicted <GPB>

F:578-594/Domain: transmembrane #status predicted <TM1>

F:770-786/Domain: transmembrane #status predicted <TM2>

F:795-811/Domain: transmembrane #status predicted <TM3>

F:48,110,164,278,421,453,505,564,692/Binding site: carbohydrate (Asn) (covalent) #status

Query Match

Best Local Similarity 41.5%; Score 49; DB 1; Length 917;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 IHHTFRGIVHAGRSIGRFLTG 22
Db 753 LHGFEEGLDVGRAVRYVLG 773

RESULT 4

AC2951

hypothetical protein Atu3209 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AC2951

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AC2951

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-83 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAL44025.1; PID:g17741585; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu3209

A:Map position: linear chromosome

Query Match

Best Local Similarity 40.7%; Score 48; DB 2; Length 83;
Matches 11; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 2 IHHTFRGIVHAGRSIGRFLTG 22

Db 10 VHH-----GAGRRIGRFLIG 24

RESULT 5

AH0280

NAD(P) transhydrogenase (AB-specific) (EC 1.6.1.2) - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 03-Jun-2002

C:Accession: AH0280

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrer, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AH0280

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-508 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC91108.1; PID:g15980299; GSPDB:GN00175

C:Genetics:

A:Gene: pntA

C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogenase

C:Keywords: oxidoreductase

Query Match

Best Local Similarity 40.7%; Score 48; DB 2; Length 508;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIVHAGRSIGRFLTG 22
Db 139 YRAIVEAHEFGRRFTG 155

RESULT 6

AB0683

pyridine nucleotide transhydrogenase chain-alpha [imported] - Salmonella enterica sub

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB0683

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AB0683

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-509 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01835.1; PID:g16502679; GSPDB:GN00176

C:Genetics:

A:Gene: STY1589

C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogena

Query Match

Best Local Similarity 40.7%; Score 48; DB 2; Length 509;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIVHAGRSIGRFLTG 22
Db 139 YRAIVEAHEFGRRFTG 155

RESULT 7

DEECXA

NAD(P) transhydrogenase (B-specific) (EC 1.6.1.1) alpha chain - Escherichia coli (str N; Alternate names: pyridine nucleotide transhydrogenase alpha chain

C:Species: Escherichia coli

```

C:/Date: 30-Jun-1988 #sequence_revision 19-May-1995 #text_change 03-Jun-2002
C:/Accession: S24380; A25012; E64916; S18355; S23833
R:/Ahmad, S.; Glavas, N.A.; Bragg, P.D.
Eur. J. Biochem. 207, 733-739, 1992
A:/Title: A mutation at G1314 of the beta subunit of the Escherichia coli pyridine nucle
A:/Reference number: S24380; MUID:92339464; PMID:1633824
A:/Accession: S24380
A:/Molecule type: DNA
A:/Residues: 1-510 <AHM>
A:/Cross-references: EMBL:X66086; NID:g42455; PIDN:CAA46884.1; PID:g42456
R:/Clarke, D.M.; Loo, T.W.; Gillam, S.; Bragg, P.D.
Eur. J. Biochem. 158, 647-653, 1986
A:/Title: Nucleotide sequence of the pntA and pntB genes encoding the pyridine nucleotide
A:/Reference number: A91172; MUID:86274751; PMID:3525165
A:/Accession: A25012
A:/Molecule type: DNA
A:/Residues: 1-39, 'VNMQVLTIRKLCG', 56-328, 'R', 330-502 <CLA>
R:/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:/Title: The complete genome sequence of Escherichia coli K-12.
A:/Reference number: A64720; MUID:97426617; PMID:9278503
A:/Accession: E64916
A:/Status: nucleic acid sequence not shown; translation not shown
A:/Molecule type: DNA
A:/Residues: 1-510 <BLAT>
A:/Cross-references: GB:AE000255; GB:U00096; NID:g1787875; PIDN:AACT4675.1; PID:g1787887;
A:/Experimental source: strain K-12, substrain MG1655
R:/Tong, R.C.W.; Glavas, N.A.; Bragg, P.D.
Biochim. Biophys. Acta 1080, 19-28, 1991
A:/Title: Topological analysis of the pyridine nucleotide transhydrogenase of Escherichia
A:/Reference number: S18355; MUID:92031646; PMID:1932078
A:/Accession: S18355
A:/Molecule type: protein
A:/Residues: 1-10; 16-25; 'XF', 'ISV', 136-138, 225, 'EGT', 229-238; 270-273, 'X', 275-278; 'AG', 279
C:/Genetics:
A:/Gene: pntA
A:/Map position: 35 min
C:/Complex: heterotetramer; two alpha and two beta chains
C:/Function:
A:/Description: catalyzes the reversible hydride ion transfer between NAD and NADP; this
ane
A:/Pathway: NAD phosphorylation and dephosphorylation
C:/Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogenase
C;/Keywords: heterotetramer; inner membrane; transmembrane protein; NAD; NADP; oxidoreduc
F:/1-509/Domain: NAD(P)+ transhydrogenase (B-specific) alpha chain homology <TBA>
F:/1-295/Domain: alanine dehydrogenase homology <ALA>
F:/167-195/Region: beta-alpha-beta NAD nucleotide-binding fold
F:/404-420/Domain: transmembrane #status predicted <TM1>
F:/429-445/Domain: transmembrane #status predicted <TM2>
F:/453-469/Domain: transmembrane #status predicted <TM3>
F:/478-494/Domain: transmembrane #status predicted <TM4>

Query Match          40.7%; Score 48; DB 1; Length 510;
Best Local Similarity 52.9%; Pred. No. 16;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      6 FRGIVHAGRSTIGRFLTG 22
      :| || | || | ||
Db      139 YRAIVEAAHFGRRFFTG 155

RESULT 8
C86729
hypothetical protein y1eh [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:/Species: Lactococcus lactis subsp. lactis
C:/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:/Accession: C86729
R:/Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A:/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:/Reference number: A86625; MUID:21235186; PMID:11337471
A:/Accession: C86729

```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-883 <STO>
A;Cross-references: GB:AE005176; PID:g12723760; PIDN:AAK04933.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: y1eH

Query	Match	40.7%;	Score	48;	DB	2;	Length	883;	
Best	Local Similarity	45.0%;	Pred. No.	29;					
Matches	9;	Conservative	4;	Mismatches	7;	Indels	0;	Gaps	0;
QY	2	IHHIFRGIVHAGRSIGRFLT	21						
		:: : :							
Db	51	IHSLYRNILHSGSSGFLYT	70						

RESULT 9
AG2840

alcohol dehydrogenase adh [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AG2840
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <KUR>
A:Cross-references: GB:AE008688; PIDN:AL43141.1; PID:q17740616; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: adh
A:Map position: circular chromosome
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match	40.3%;	Score 47.5;	DB 2;	Length 347;
Best Local Similarity	50.0%;	Pred. No. 13;		
Matches 10;	Conservative 4;	Mismatches 5;	Indels 1;	Gaps 1;

QY 4 HIFGI VHA-GRSIGRFLTG 22
| | : : | | : : | | | |
Db 60 HEFAGVISAVGRNVTREKFG 79

RESULT 10
A97618

alcohol dehydrogenase (adh-ht) [imported] - Agrobacterium tumefaciens (strain C58, Ce
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: A97618
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A/Reference number: A97359; PMID:11743194

A;Accession: A97618
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK87898.1; PID:g15157292; GSPDB:GN00169
C:Genetics:

Query Match	Score	DB 2	Length	357;
40.38;	Score 47.5;			

A:Accession: T11067
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-360 <KED>
A:Cross-references: EMBL:AF015193; NID:g2735934; PID:g2735939; PIDN:AAC61614.1
A:Experimental source: strain forest
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:1-196/Domain: cytochrome b6 homology <CB6>
F:70,169/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:84,183/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 39.0%; Score 46; DB 2; Length 360;
Best Local Similarity 45.8%; Pred. No. 23;
Matches 11; Conservative 7; Mismatches 4; Indels 2; Gaps 2;

QY 1 FIH-HIFRGLVHAG-RSIGRFLTG 22
||:||||:|:::| ||:|:|
Db 80 FIYLIHFKGLTYGSYRLIGVWLSG 103

Search completed: May 17, 2003, 14:28:01
Job time : 45 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:15:39 ; Search time 25 Seconds
(without alignments)
36.499 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FIHHIFRGIVHAGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.5	41.5	917	1	VGLB_HSVB2 P12641 bovine herp
2	48	40.7	510	1	PNTA_ECOLI P07001 escherichia
3	45.5	38.6	388	1	CYB_MAIZE P04165 zea mays (m
4	45.5	38.6	392	1	CYB_SOLITU P29757 solanum tub
5	45.5	38.6	392	1	CYB_VICFA P05718 vicia faba
6	45.5	38.6	394	1	CYB_OENBE P09843 oenothera b
7	45.5	38.6	397	1	CYB_ORYSA P14833 oryza sativ
8	45	38.1	666	1	PD14_MOUSE Q92183 mus musculu
9	44	37.3	336	1	SYRM_RHIEF O08812 rhizobium e
10	44	37.3	380	1	SYW_HALNI Q9hn66 halobacteri
11	44	37.3	776	1	PSAB_HETTR Q9xqv2 heterocapsa
12	43	36.4	395	1	ARP2_CAEEL P33489 caenorhabdi
13	43	36.4	445	1	NSMA_CAEEL O45870 halobacteri
14	43	36.4	492	1	GVD2_HALNI Q9hnt2 halobacteri
15	43	36.4	492	1	GVD2_HALNI P33958 halobacteri
16	43	36.4	663	1	PD15_HUMAN Q9um07 homo sapien
17	43	36.4	1436	1	DPO3_STAMM Q53665 staphylococ
18	43	36.4	1438	1	DPO3_STAMM Q99uk8 staphylococ
19	42.5	36.0	515	1	CXAA_HUMAN P57773 homo sapien
20	42	35.6	354	1	SXL_DROSTU Q24668 drosophila
21	42	35.6	362	1	P2YR_CHICK P34996 gallus gall
22	42	35.6	362	1	P2YR_MELGA P49652 meleagris g
23	42	35.6	378	1	H182_CAUCR Q9a5b6 caulobacter
24	42	35.6	666	1	PD14_RAT O88807 rattus norv
25	41.5	35.2	467	1	LEU2_VIBCH Q9kp81 vibrio chol
26	41	34.7	114	1	KDGL_BACSU P19638 bacillus su
27	41	34.7	201	1	GCHI_STRCO Q9x813 streptomyc
28	41	34.7	270	1	TRUA_ECOLI P07649 escherichia
29	41	34.7	411	1	EF1G_TRYCR P34715 trypanosoma
30	41	34.7	521	1	NU2C_SYNP7 P29801 synecchococ
31	41	34.7	586	1	HO_YEAST P09932 saccharomyc
32	41	34.7	885	1	VGLB_HSV2S P24994 herpes simp
33	40.5	34.3	170	1	OBPB_HUMAN Q9nph6 homo sapien

34	40.5	34.3	250	1	RK13_SPIOL	P12629 spinacia ol
35	40	33.9	181	1	NDKM_COLLI	P87355 columba liv
36	40	33.9	219	1	CLNI_HUMAN	Q9bxu9 homo sapien
37	40	33.9	219	1	CLNI_MOUSE	Q9j197 mus musculu
38	40	33.9	225	1	CYPD_YEAST	P35176 saccharomyc
39	40	33.9	322	1	GDC_RAT	P16261 rattus norv
40	40	33.9	330	1	GDC_BOVIN	Q01888 bos taurus
41	40	33.9	425	1	Y461_MYCGE	P47699 mycoplasma
42	40	33.9	425	1	Y461_MYCPN	P75115 mycoplasma
43	40	33.9	512	1	PNTA_HAEIN	P43842 haemophilus
44	40	33.9	676	1	NTP2_VACCA	O57193 vaccinia vi
45	40	33.9	676	1	NTP2_VACCC	P20502 vaccinia vi

ALIGNMENTS

RESULT 1						
VGLB_HSVB2						
ID	VGLB_HSVB2	STANDARD:	PRT:	917	AA.	
AC	P12641:					
DT	01-OCT-1989 (Rel. 12, Created)					
DT	01-AUG-1990 (Rel. 15, Last sequence update)					
DT	16-OCT-2001 (Rel. 40, Last annotation update)					
DE	Glycoprotein B-1 precursor.					
OS	Bovine herpesvirus type 2 (strain BMV) (Bovine mammillitis virus).					
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;					
OC	Alphaherpesvirinae; Simplexvirus.					
OX	NCBI_TaxID=10296;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88306231; PubMed=2841793;					
RA	Hammerschmidt W., Contraths F., Mankertz J., Pauli G., Ludwig H.,					
RA	Buhk H.-J.;					
RT	"Conservation of a gene cluster including glycoprotein B in bovine					
RT	herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";					
RL	Virology 165:388-405(1988).					
RN	[2]					
RP	SEQUENCE OF 1-200 FROM N.A.					
RX	MEDLINE=88306232; PubMed=2457278;					
RA	Hammerschmidt W., Contraths F., Mankertz J., Buhk H.-J., Pauli G.,					
RA	Ludwig H.;					
RT	"Common epitopes of glycoprotein B map within the major DNA-binding					
RT	proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex					
RT	virus type 1 (HSV-1).";					
RL	Virology 165:406-418(1988).					
CC	-!- FUNCTION: GB1 IS A 130 kDa GLYCOPROTEIN WHICH IS NECESSARY FOR THE					
CC	PENETRATION OF THE VIRUS INTO THE HOST CELL AND THE INDUCTION OF A					
CC	SYNCYTIAL PHENOTYPE.					
CC	-!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; M21628; AAA46053.1; -					
DR	EMBL; M21632; AAA46052.1; -					
DR	PIR; C29242; VGBEBH.					
DR	InterPro; IPR000234; Glycoprot_B.					
DR	Pfam; PF00606; Glycoprotein_B; 1.					
DR	Prodom; PD000693; Glycoprot_B; 1.					
KW	Glycoprotein; Transmembrane; Signal.					
FT	SIGNAL	1	22			
FT	CHAIN	23	917			GLYCOPROTEIN B-1.
FT	TRANSMEM	578	594			POTENTIAL.
FT	TRANSMEM	770	786			POTENTIAL.
FT	TRANSMEM	795	811			POTENTIAL.
FT	CARBOHYD	48	48			N-LINKED (GLCNAC: . .) (POTENTIAL).
FT	CARBOHYD	110	110			N-LINKED (GLCNAC: . .) (POTENTIAL).

```

FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 917 AA; 101882 MW; 1B96CBF50DB4D3F3 CRC64;

Query Match 41.5%; Score 49; DB 1; Length 917;
Best Local Similarity 38.1%; Pred. No. 8.1;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGRSIGRFLTG 22
Db 753 LHGFEGGLGVDGVRVVLG 773

RESULT 2
PNTA_ECOLI STANDARD; PRT; 510 AA.
ID PNTA_ECOLI
AC P07001; P76888;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NAD(P) transhydrogenase subunit alpha (EC 1.6.1.2) (Pyridine
DE nucleotide transhydrogenase subunit alpha) (Nicotinamide nucleotide
DE transhydrogenase subunit alpha).
GN PNTA OR B1603.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86274751; Pubmed=3525165;
RA Clarke D.M., Loo T.W., Gillam S., Bragg P.D.;
RT "Nucleotide sequence of the pntA and pntB genes encoding the pyridine
RT nucleotide transhydrogenase of Escherichia coli.";
RL Eur. J. Biochem. 158:647-653(1986).
RN [2]
RP REVISIONS.
RX MEDLINE=92339464; Pubmed=1633824;
RA Ahmad S., Glavas N.A., Bragg P.D.;
RT "A mutation at Gly314 of the beta subunit of the Escherichia coli
RT pyridine nucleotide transhydrogenase abolishes activity and affects
RT the NADP(H)-induced conformational change.";
RL Eur. J. Biochem. 207:733-739(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; Pubmed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; Pubmed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horluchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -!- FUNCTION: THE TRANSHYDROGENATION BETWEEN NADH AND NADP IS COUPLED
```

```

CC TO RESPIRATION AND ATP HYDROLYSIS AND FUNCTIONS AS A PROTON PUMP
CC ACROSS THE MEMBRANE.
CC -!- CATALYTIC ACTIVITY: NADPH + NAD(+) = NADP(+) + NADH.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: TO THE N-TERMINAL PART OF THE EUKARYOTIC MITOCHONDRIAL
CC ENZYME AND TO ALANINE DEHYDROGENASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04195; CAB37089.1; -.
DR EMBL; X66086; CAA46884.1; -.
DR EMBL; AE000255; AAC74675.1; -.
DR EMBL; D90803; BAA15342.1; -.
DR EMBL; D90804; BAA15352.1; -.
DR EMBL; D90802; BAA15337.1; -.
DR PIR; A25012; DEECXA.
DR PIR; S24380; S24380.
DR PIR; S23833; S23833.
DR HSSP; Q60164; 1HZZ.
DR Ecogene; EG10744; pntA.
DR InterPro; IPR004002; Aladh_PNT.
DR InterPro; IPR004571; pntA.
DR Pfam; PF01262; Aladh_PNT; 1.
DR TIGRfams; TIGR00561; pntA; 1.
DR PROSITE; PS00836; ALADH_PNT_1; 1.
DR PROSITE; PS00837; ALADH_PNT_2; 1.
DR Oxidoreductase; NAD; NADP; Transmembrane; Inner membrane;
KW Complete proteome.
FT NP_BIND 167 197 NAD (BY SIMILARITY).
FT TRANSMEM 402 422 POTENTIAL.
FT TRANSMEM 423 443 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
SQ SEQUENCE 510 AA; 54623 MW; 801742097BEA6943 CRC64;

Query Match 40.7%; Score 48; DB 1; Length 510;
Best Local Similarity 52.9%; Pred. No. 6.5;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGLVHAGRSIGRFLTG 22
Db 139 YRAIVEAHEFGREFTG 155

RESULT 3
CYB_MAIZE STANDARD; PRT; 388 AA.
ID CYB_MAIZE
AC P04165;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Zea mays (Maize).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Dawson A.J., Jones V.P., Leaver C.J.;
RT "The apocytochrome b gene in maize mitochondria does not contain
RT introns and is preceded by a potential ribosome binding site.";
RL EMBO J. 3:2107-2113(1984).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
```

```
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -I- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -I- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X00789; CAA25367.1; -
CC PIR; A00156; CB2M.
CC -----
CC MaizedB; 69227; -
CC InterPro; IPR000179; Cyt_b_b6.
CC Pfam; PF00032; cytochrome_b_c1.
CC Pfam; PF00033; cytochrome_b_n; 1.
CC PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
CC PROSITE; PS00193; CYTOCHROME_B_OO; 1.
CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 88 88 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 102 102 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 189 189 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 203 203 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 388 AA; 43567 MW; 56CF734C2B62B3F6 CRC64;

Query Match 38.6%; Score 45.5; DB 1; Length 388;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 IH-HIFRGIVHAGRSIGR 18
Db 99 VHLHIFRGLYHASYSR 116

RESULT 4
CYB_SOLTU STANDARD: PRT; 392 AA.
ID CYB_SOLTU
AC P29757;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Solanum tuberosum (Potato).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bintje;
RX MEDLINE=91355947; PubMed=1840690;
RA Zanolungo S., Litvak S., Jordana X.;
RT "Isolation and nucleotide sequence of the potato mitochondrial gene
RL for apocytochrome b.";
RL Plant Mol. Biol. 17:527-530(1991).

[2]
RP SEQUENCE OF 1-18.
RC TISSUE=Tuber;
RX MEDLINE=93131029; PubMed=8420797;
RA Braun H.-P., Schmitz U.K.;
RT "Purification and sequencing of cytochrome b from potato reveals
RL methionine cleavage of a mitochondrially encoded protein.";
RL FEBS Lett. 316:128-132(1993).
```

```
CC -I- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -I- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -I- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -I- PTM: FIRST MITOCHONDRIAL ENCODED PROTEIN TO BE SHOWN TO HAVE ITS
CC N-TERMINAL METHIONINE CLEAVED OFF.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X58437; CAA41343.1; -
CC PIR; S17427; CBPOM.
CC InterPro; IPR000179; Cyt_b_b6.
CC Pfam; PF00032; cytochrome_b_c1.
CC Pfam; PF00033; cytochrome_b_n; 1.
CC PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
CC PROSITE; PS00193; CYTOCHROME_B_OO; 1.
CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT INIT_MET 0 0
FT METAL 87 87 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 101 101 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 188 188 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 202 202 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 392 AA; 43913 MW; CA147084333F73AC CRC64;

Query Match 38.6%; Score 45.5; DB 1; Length 392;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 IH-HIFRGIVHAGRSIGR 18
Db 98 VHLHIFRGLYHASYSR 115

RESULT 5
CYB_VICFA STANDARD: PRT; 392 AA.
ID CYB_VICFA
AC P05718;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Vicia faba (Broad bean).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxID=3906;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88303319; PubMed=3405753;
RA Wahleithner J.A., Wolstenholme D.R.;
RT "Ribosomal protein S14 genes in broad bean mitochondrial DNA.";
RL Nucleic Acids Res. 16:6897-6913(1988).
CC -I- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -I- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -I- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
```

```

CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07237; CAA30226.1; -.
DR PIR: S01221; CBVF.
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam: PF00032; cytochrome_b_c; 1.
DR Pfam: PF00033; cytochrome_b_n; 1.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
KW METAL 88 88 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 102 102 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 189 189 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 203 203 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 392 AA; 44013.MW; 023D3396B2FA2BD3 CRC64;

```

Query Match	38.6%;	Score 45.5;	DB 1;	Length 392;
Best Local Similarity	55.6%;	Pred. No. 12;		
Matches 10;	Conservative 2;	Mismatches 5;	Indels 1;	Gaps 1;

```
QY      2 IH-HIRGIVHAGRSIGR 18
          :| | | | | : | | |
Db      99 VHLHIFRGLYHASYSSPR 116
```

```

CC      329 AND 363 ARE MODIFIED BY RNA EDITING.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X07126; CAA30135.1; -.
DR      EMBL; X17030; CAA34892.1; -.
DR      PIR; S20141; CBOBE.
DR      InterPro; IPR000179; Cyt_b_b6.
DR      Pfam; PF00032; cytochrome_b_C; 1.
DR      Pfam; PF00033; cytochrome_b_N; 1.
DR      PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR      PROSITE; PS00193; CYTOCHROME_B_OO; 1.
KW      Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW      Heme; RNA editing.
FT      METAL      89      89      IRON 1 (HEME B562 AXIAL LIGAND).
FT      METAL      103     103     IRON 2 (HEME B566 AXIAL LIGAND).
FT      METAL      191     191     IRON 2 (HEME B562 AXIAL LIGAND).
FT      METAL      204     204     IRON 1 (HEME B566 AXIAL LIGAND).
SQ      SEQUENCE   394 AA;  44152 MW;  D8CAEB874886C20FA CRC64;

```

Query Match	38.6%	Score 45.5;	DB 1;	Length 394;
Best Local Similarity	55.6%;	Pred. No. 12;		
Matches 10; Conservative	2;	Mismatches 5;	Indels 1;	Gaps 1;

QY 2 IH-HIFRGIVHAGRSIGR 18
: | | | | : | | |
Db 100 VHLHIFRGLYHASYSPPR 117

RESULT	6
CYB_OENBE	
ID	CYB_OENBE
AC	P09843;
DT	01-MAR-1989 (Rel. 10, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Cytochrome B.
GN	COB OR CYTB.
OS	Oenothera bertiana (Bertero's evening primrose).
OG	Mitochondrion.
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Myrtales; Onagraceae; Oenothera.
NCBI_TaxID	3950;
RN	[1]

RP SEQUENCE FROM N.A.
 RA Schuster W., Brennick A.;
 RT "TGA-Termination codon in the apocytochrome b gene from Oenothera
 RL mitochondria.";
 RN Curr. Genet. 9:157-163(1985).
 [2]
 RP SEQUENCE OF 1-21 FROM N.A., AND RNA EDITING.
 RX MEDLINE=90221801; PubMed=2326162;
 RA Schuster W., Unsold M., Wissing B., Brennick A.;
 RT "Ribosomal protein S14 transcripts are edited in Oenothera
 RL mitochondria.";
 RNL Nucleic Acids Res. 18:229-233(1990).
 CCC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CCC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CCC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CCC COUPLED TO ATP SYNTHESIS.
 CCC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CCC BOUND TO THE PROTEIN.
 CCC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CCC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CCC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CCC -1- CAUTION: POSITIONS 19, 97, 101, 104, 110, 121, 137, 191, 286, 304,

CC 329 AND 363 ARE MODIFIED BY RNA EDITING.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X07126; CAA30135.1; - .
DR EMBL; X17030; CAA34892.1; - .
DR PIR; S20141; CBOEF.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_C; 1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme; RNA editing.

FT	METAL	89	89	IRON 1 (HEME B562 AXIAL LIGAND).
FT	METAL	103	103	IRON 2 (HEME B566 AXIAL LIGAND).
FT	METAL	191	191	IRON 2 (HEME B562 AXIAL LIGAND).
FT	METAL	204	204	IRON 1 (HEME B566 AXIAL LIGAND).
SQ	SEQUENCE	394 AA;	44152 MW;	D8CAE874886C20FA CRC64;

Query Match	38.6%	Score 45.5;	DB 1;	Length 394;
Best Local Similarity	55.6%;	Pred. No. 12;		
Matches 10; Conservative	2;	Mismatches 5;	Indels 1;	Gaps 1;

QY 2 IH-HIFRGIVHAGRSIGR 18
: | | | | : | | |
Db 100 VHLHIFRGLYHASYSPPR 117

RESULT 7		
CYB_ORYSA		
ID	CYB_ORYSA	STANDARD; PRT; 397 AA.
AC	P14833;	
DT	01-APR-1990 (Rel. 14, Created)	
DT	01-APR-1990 (Rel. 14, Last sequence update)	
DT	30-MAY-2000 (Rel. 39, Last annotation update)	
DE	Cytochrome B.	
GN	COB OR CYTB.	
OS	Oryza sativa (Rice).	
OG	Mitochondrion.	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC	Ehrhartoideae; Oryzeae; Oryza.	
OX	NCBI_TaxID=4530;	

SEQUENCE FROM N.A.
STRAIN=cv. Indica-IR36;
MEDLINE=90221830; PubMed=2326174;
Kaleikau E.K., Andre C.P., Doshi B., Walbot V.;
"Sequence of the rice mitochondrial gene for apocytochrome b.";
Nucleic Acids Res. 18:372-372(1990).
-!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
COUPLED TO ATP SYNTHESIS.
-!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
BOUND TO THE PROTEIN.
-!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CYTOCHROME C1 AND THE RIESKE PROTEIN.
-!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)


```
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X17064; CAA34910.1; -.
DR      PIR: JQ0164; CBR2.
DR      InterPro: IPR000179; Cyt_b_b6.
DR      Pfam: PF00032; cytochrome_b_c; 1.
DR      Pfam: PF00033; cytochrome_b_n; 1.
DR      PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR      PROSITE: PS00193; CYTOCHROME_B_OO; 1.
KW      Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
Heme.
FT      METAL      88      88      IRON 1 (HEME B562 AXIAL LIGAND).
FT      METAL      102     102     IRON 2 (HEME B566 AXIAL LIGAND).
FT      METAL      189     189     IRON 2 (HEME B562 AXIAL LIGAND).
FT      METAL      203     203     IRON 1 (HEME B566 AXIAL LIGAND).
SQ      SEQUENCE      397 AA; 44540 MW; D9122B9F9C20D6DE CRC64;

Query Match
Best Local Similarity      38.6%; Score 45.5; DB 1; Length 397;
Matches      10; Conservative      2; Mismatches      5; Indels      1; Gaps      1;

OY      2 IH-HIFRGIVHAGRSIGR 18
DB      99 VHLHIFRGIVHAGRSYSSPR 116

RESULT 8
PD14_MOUSE      STANDARD;      PRT;      666 AA.
AC      Q92183;
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Protein-arginine deiminase type IV (EC 3.5.3.15) (Peptidylarginine
DE      deiminase IV).
GN      PAD14 OR PD14.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Epidermis;
RX      MEDLINE=99192810; PubMed=10092850;
RA      Rusc A.A., Ikejiri Y., Ono H., Yonekawa T., Shiraiwa M., Kawada A.,
RA      Takahara H.;
RT      "Molecular cloning of cDNAs of mouse peptidylarginine deiminase type
RT      I, type III and type IV, and the expression pattern of type I in
RT      mouse.";
RL      Eur. J. Biochem. 259:660-669(1999).
CC      -!- FUNCTION: CATALYZES THE DEIMINATION OF ARGININE RESIDUES OF
CC      PROTEINS.
CC      -!- CATALYTIC ACTIVITY: Protein L-arginine + H(2)O = protein L-
CC      citrulline + NH(3).
CC      -!- COFACTOR: REQUIRES CALCIUM IONS.
CC      -!- SIMILARITY: SOME SIMILARITY WITH ARGINASE AND TO ARGININE
CC      DEIMINASE.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AB013850; BAA34246.1; -.
DR      MGD; MGI:1338898; PD14.
DR      InterPro: IPR004303; Prot_arg_deim.
DR      Pfam: PF03068; PAD; 1.
KW      Hydrolase; Calcium-binding; Multigene family.
DT      CA_BIND      505      516      EF-HAND (POTENTIAL).
DT      SEQUENCE      666 AA; 74476 MW; 70FAE4E7E232D34A CRC64;
```

```
Query Match
Best Local Similarity      38.1%; Score 45; DB 1; Length 666;
Matches      6; Conservative      6; Mismatches      6; Indels      0; Gaps      0;

OY      1 FIHHIFRGIVHAGRSIGR 18
DB      637 YTYHYNGEVHCGTIVRR 654

RESULT 9
SYRM_RHIEF      STANDARD;      PRT;      336 AA.
ID      SYRM_RHIEF
AC      Q08812;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Syrm protein (Symbiotic regulator).
GN      SYRM.
OS      Rhizobium etli.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Rhizobiaceae; Rhizobium.
OX      NCBI_TaxID=29449;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CNPAF512;
RX      MEDLINE=93376517; PubMed=8367305;
RA      Michiels J., de Wilde P., Vanderleyden J.;
RT      "Sequence of the Rhizobium leguminosarum biovar phaseoli syrm gene.";
RL      Nucleic Acids Res. 21:3893-3893(1993).
CC      -!- FUNCTION: SYRM ACTS IN TRANS TO STIMULATE NOD GENE EXPRESSION VIA
CC      NODD3 AND EXO GENE EXPRESSION VIA SYRA.
CC      -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC      REGULATORS.
CC      -!- CAUTION: STRAIN CNPAF512 WAS ORIGINALLY THOUGHT TO ORIGINATE FROM
CC      R.LEGUMINOSARUM BV. PHASEOLI.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Z23013; CAA80556.1; -.
DR      InterPro: IPR000847; HTH_LYSR.
DR      InterPro: IPR005119; LYSR_subst.
DR      Pfam: PF00126; HTH_1; 1.
DR      Pfam: PF03466; LYSR_substrate; 1.
DR      PROSITE: PS00044; HTH_LYSR_FAMILY; 1.
KW      Modulation; Transcription regulation; DNA-binding; Activator.
FT      DNA_BIND      58      77      H-T-H MOTIF (BY SIMILARITY).
FT      SEQUENCE      336 AA; 37800 MW; FACF9463A6A872BD CRC64;

Query Match
Best Local Similarity      37.3%; Score 44; DB 1; Length 336;
Matches      7; Conservative      3; Mismatches      3; Indels      0; Gaps      0;

OY      6 FRGIVHAGRSIGR 18
DB      55 YRNVTHAGQHIGR 67

RESULT 10
SYW_HALN1      STANDARD;      PRT;      380 AA.
ID      SYW_HALN1
AC      Q9HN66;
DT      15-JUN-2002 (Rel. 41, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Tryptophanyl-trna synthetase (EC 6.1.1.2) (Tryptophan--trna ligase)
DE      (Trprs).
```

```
GN TRPS OR TRPS2 OR VNG2232G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005109; AAG20355.1; -
DR InterPro; IPR002306; TRP-trna-synt_1b.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMS; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA-trna_ligase_1; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 74 82 "HIGH" REGION.
FT SITE 249 253 "KMSKS" REGION.
SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;
Query Match 37.3%; Score 44; DB 1; Length 380;
Best Local Similarity 53.3%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 7 RGIHVAGRSIGRFLT 21
DB 48 RGLFYAGRDVDFLT 62
RESULT 11
PSAB_HETTR STANDARD; PRT; 776 AA.
AC Q9XQV2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PSI-B).
GN PSAB.
OS Heterocapsa triquetra (Dinoflagellate).
OG Chloroplast.
OC Eukaryota; Alveolata; Dinophyceae; Peridinales; Heterocapsaceae;
OC Heterocapsa.
OX NCBI_TaxID=66468;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCMP 449;
RX MEDLINE=99334925; PubMed=10408440;
RA Zhang Z., Green B.R., Cavalier-Smith T.;
```

```
RT "Single gene circles in dinoflagellate chloroplast genomes."
RL Nature 400:155-159(1999).
CC -1- FUNCTION: Psaa and psab bind P700, the primary electron donor of
CC photosystem I (PSI), as well as the electron acceptors A0, A1, and
CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin
CC oxidoreductase.
CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1
CC is a phylloquinone and FX is a 4Fe-4S iron-sulfur center.
CC -1- SUBUNIT: A psaa/B heterodimer binds the P700 chlorophyll special
CC pair and subsequent electron acceptors. The PSI reaction center of
CC higher plants and algae is composed of one at least 11 subunits.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
CC thylakoid membrane.
CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF130032; AAD44699.1; -
DR HSSP; P25897; 1JB0.
DR InterPro; IPR001280; PSI_Psaa/B.
DR Pfam; PF00223; psaa_psab; 1.
DR PRINTS; PR00257; PHOTOSYPSAAB.
DR PROSITE; PS00419; PHOTOSYSTEM_I_PSAAB; 1.
KW Photosynthesis; Photosystem I; Transport; Electron transport;
KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;
KW Iron-sulfur; 4Fe-4S; Chlorophyll.
FT TRANSMEM 57 80 I (POTENTIAL).
FT TRANSMEM 150 172 II (POTENTIAL).
FT TRANSMEM 233 257 III (POTENTIAL).
FT TRANSMEM 334 352 IV (POTENTIAL).
FT TRANSMEM 382 405 V (POTENTIAL).
FT TRANSMEM 421 447 VI (POTENTIAL).
FT TRANSMEM 472 494 VII (POTENTIAL).
FT TRANSMEM 561 579 VIII (POTENTIAL).
FT TRANSMEM 619 640 IX (POTENTIAL).
FT TRANSMEM 687 709 X (POTENTIAL).
FT TRANSMEM 753 773 XI (POTENTIAL).
FT TRANSMEM 603 603 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 612 612 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 698 698 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
FT BINDING 698 698 LIGAND (BY SIMILARITY).
FT BINDING 706 706 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 714 714 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 715 715 A1 PHYLLOQUINONE (BY SIMILARITY).
FT BINDING 746 746 A1 PHYLLOQUINONE (BY SIMILARITY).
SQ SEQUENCE 776 AA; 86143 MW; F8640CA5F58C8913 CRC64;
Query Match 37.3%; Score 44; DB 1; Length 776;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 3 HHIFRGIHVAGRSIGRFLTG 22
DB 302 NHIFGSTVGAGKAILTF LG 321
RESULT 12
ARP2_CAEEL STANDARD; PRT; 395 AA.
ID ARP2_CAEEL
AC P53489;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Actin-like protein 2 (Actin-like protein C).
GN ACTC OR K07C5.1.
OS Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
```



```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurtry A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PART OF A COMPLEX IMPLICATED IN THE CONTROL OF ACTIN
CC POLYMERIZATION IN CELLS (BY SIMILARITY).
CC -!- SUBUNIT: BELONGS TO A COMPLEX COMPOSED OF ARP2, ARP3, P41-ARC,
CC P34-ARC, P21-ARC, P20-ARC AND P16-ARC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE ACTIN FAMILY. ARP2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z71181; CA94894.1; -.
DR WormPep; K07C5.1; CE06111.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; actin; 1.
DR SMART; SM00268; ACTIN; 1.
DR PROSITE; PS00432; ACTINS_2; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
KW Structural protein; Cytoskeleton.
SQ SEQUENCE 395 AA; 44840 MW; D13B735000420692 CRC64;

Query Match
Best Local Similarity 36.4%; Score 43; DB 1; Length 395;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGRSIGRFL 20
Db 175 LHLLTRRLDIAGRDITKYL 193

RESULT 13
NSMA_CAEEL STANDARD; PRT; 445 AA.
AC 045870;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative neutral sphingomyelinase (EC 3.1.4.12).
GN T27F6.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Dobson R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Sphingomyelin + H(2)O = N-acylsphingosine +
CC choline phosphate.
CC -!- SIMILARITY: BELONGS TO THE NEUTRAL SPHINGOMYELINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z82060; CAB04885.1; -.
DR WormPep; T27F6.6; CE16515.

```

```

DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
KW Hypothetical protein; Hydrolase; Transmembrane.
FT TRANSMEM 377 399
FT TRANSMEM 377 399
FT METAL 403 425
FT METAL 83 83
FT SITE 215 215
FT ACT SITE 318 318
SQ SEQUENCE 445 AA; 51054 MW; 969FEC23A037FF72 CRC64;

Query Match
Best Local Similarity 47.1%; Score 43; DB 1; Length 445;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FIHIFRGIVHAGRSIG 17
Db 137 FAHHIHRGDMFGCKVVG 153

RESULT 14
GVD2_HALN1 STANDARD; PRT; 492 AA.
ID GVD2_HALN1
AC Q9HHT2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GvpD protein 2.
GN GVPD2 OR VNG6240G.
OS Halobacterium sp. (strain NRC-1).
OC Plasmid pNRC200.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: MAY PLAY A ROLE IN AN ENERGY REQUIRING PROCESS SUCH AS
CC ASSEMBLY OF GAS VESICLES IN ADDITION TO A POSSIBLE STRUCTURAL OR
CC REGULATORY FUNCTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE005156; AAG20894.1; -.
KW Gas vesicle; ATP-binding; Plasmid; Complete proteome.
FT NP_BIND 39 46
FT ATP (POTENTIAL).
SQ SEQUENCE 492 AA; 53928 MW; 655965FD1C36331B CRC64;

Query Match
Best Local Similarity 36.4%; Score 43; DB 1; Length 492;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 8 GIVHAGRSIGRFLTG 22
Db 15 GIAHPREIRRFETG 29

```

RESULT 15
GVD2_HALN2 STANDARD; PRT; 492 AA.
ID GVD2_HALN2
AC P33958;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GVPD protein 2.
GN GVPD2 OR GVPD.
OS Halobacterium sp. (strain NRC-817).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=148370;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021102; PubMed=1404376;
RA Engleert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312339; PubMed=8763925;
RA Krueger K., Pfeifer F.;
RT "Transcript analysis of the c-vac region and differential synthesis of
RT the two regulatory gas vesicle proteins GvpD and GvpE in Halobacterium
RT salinarum PH4.";
RL J. Bacteriol. 178:4012-4019(1996).
CC -!- FUNCTION: MAY PLAY A ROLE IN AN ENERGY REQUIRING PROCESS SUCH AS
CC ASSEMBLY OF GAS VESICLES IN ADDITION TO A POSSIBLE STRUCTURAL OR
CC REGULATORY FUNCTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X64730; CAA45992.1; -.
DR EMBL: X94688; CAA64343.1; -.
DR PIR: S28139; S28139.
KW Gas vesicle; ATP-binding.
FT NP_BIND 39 46 ATP (POTENTIAL).
SQ SEQUENCE 492 AA; 54100 MW; 3D30F6E317ABA34E CRC64;
Query Match 36.4%; Score 43; DB 1; Length 492;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Search completed: May 17, 2003, 14:25:41
Job time : 27 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:21:19 ; Search time 83 Seconds
(without alignments)
54.615 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FIHHIFRGIVHAGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SP archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriaph:*
 - 17: sp_archaeap:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	73.7	79	13 Q8UUG0	Q8ung0 morone saxa
2	84	71.2	79	13 Q8UUG2	Q8ung2 morone chry
3	53	44.9	523	2 Q9ALA2	Q9ala2 rhodobacter
4	53	44.9	524	16 Q8UCJ1	Q8ucj1 agrobacteri
5	49	41.5	205	16 Q985H7	Q985h7 rhizobium l
6	49	41.5	763	10 Q9FR59	Q9fr59 picea abies
7	48	40.7	83	16 Q8UB08	Q8ub08 agrobacteri
8	48	40.7	508	16 Q8ZE79	Q8ze79 yersinia pe
9	48	40.7	509	16 Q8ZPL0	Q8zpl0 salmonella
10	48	40.7	509	16 Q8Z6X5	Q8z6x5 salmonella
11	48	40.7	883	16 Q9CHA3	Q9cha3 lactococcus
12	48	40.7	887	12 Q9YW93	Q9yw93 macropodid
13	48	40.7	894	12 Q9YW92	Q9yw92 macropodid
14	47.5	40.3	357	16 Q8UDH4	Q8udh4 agrobacteri
15	47	39.8	273	2 Q66090	Q66090 lactococcus
16	47	39.8	512	16 Q9KM26	Q9km26 vibrio chol

17	47	39.8	513	16 Q9JZNO	Q9jzn0 neisseria m
18	47	39.8	513	16 Q9JUR3	Q9jur3 neisseria m
19	47	39.8	1762	2 Q52546	Q52546 amycolatops
20	47	39.8	1763	2 Q52790	Q52790 amycolatops
21	46	39.0	360	8 Q47575	Q47575 onchocerca
22	46	39.0	480	5 Q9VJ45	Q9vj45 drosophila
23	46	39.0	517	5 Q8T063	Q8t063 drosophila
24	46	39.0	530	16 P73496	P73496 synecocyst
25	46	39.0	537	10 Q9FW78	Q9fw78 oryza sativ
26	46	39.0	579	2 Q67993	Q67993 pseudomonas
27	46	39.0	687	10 Q9LMT8	Q9lmt8 arabidopsis
28	46	39.0	767	17 Q9HID7	Q9hid7 thermoplasm
29	46	39.0	880	16 Q8YXN7	Q8yxn7 anabaena sp
30	45.5	38.6	212	8 Q9T5A5	Q9t5a5 solanum tub
31	45.5	38.6	251	16 Q8XW76	Q8xw76 ralstonia s
32	45.5	38.6	291	8 Q35221	Q35221 oenothera b
33	45.5	38.6	392	8 Q9Z2T8	Q9zzt8 pisum sativ
34	45.5	38.6	393	8 Q95871	Q95871 nicotiana t
35	45.5	38.6	393	8 Q9MFR8	Q9mfr8 beta vulgar
36	45.5	38.6	397	8 Q35293	Q35293 oryza sativ
37	45.5	38.6	424	8 Q9B3G0	Q9b3g0 triticum ae
38	45.5	38.6	2731	11 Q9WTS4	Q9wts4 mus musculu
39	45	38.1	140	5 Q25277	Q25277 leishmania
40	45	38.1	141	2 Q31022	Q31022 vibrio chol
41	45	38.1	141	16 Q9KMH2	Q9kmh2 vibrio chol
42	45	38.1	182	13 Q90YG5	Q90yg5 oncorhynch
43	45	38.1	260	17 Q8TUN6	Q8tun6 methanosarc
44	45	38.1	321	17 Q973G7	Q973g7 sulfolobus
45	45	38.1	388	17 Q96XU4	Q96xu4 sulfolobus

ALIGNMENTS

RESULT 1				
ID	Q8UUG0	PRELIMINARY;	PRT;	79 AA.
AC	Q8UUG0;			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Moronecidin.			
OS	Morone saxatilis (Striped bass).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;			
OC	Moronidae; Morone.			
OX	NCBI_TaxID=34816;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Lauth X., Shike H., Burns J.C., Westerman M.E., Ostland V.E.,			
RA	Carlborg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shimizu C.,			
RA	Bulet P.;			
RT	"Discovery and characterization of two isoforms of moronecidin, a			
RT	novel antimicrobial peptide from hybrid striped bass.";			
RL	J. Biol. Chem. 0:0-0(2002).			
DR	EMBL: AF394244; AAL57319.1; -			
DR	EMBL: AF385583; AAL49496.1; -			
SO	SEQUENCE 79 AA; 9222 MW; C9DF13E1D4FA7EFD CRC64;			
Query Match				
	Best Local Similarity	73.7%;	Score 87;	DB 13; Length 79;
	Matches 15; Conservative	3;	Mismatches 4;	Indels 0; Gaps 0;
OY	1 FIHHIFRGIVHAGRSIGRFLTG 22			
Db	23 FFHHIFRGIVHAGKTHRLVLTG 44			
RESULT 2				
ID	Q8UUG2	PRELIMINARY;	PRT;	79 AA.
AC	Q8UUG2;			

DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Moronecidin prepropeptide precursor.
OS Morone chrysops (white bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Moronidae; Morone.
OX NCBI_TaxID=46259;
RN [1]
RP SEQUENCE FROM N.A.
RA Lauth X.C., Shike H., Burns J.C., Westerman M.E., Ostland V.E.,
RA Carlberg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shimizu C.,
RA Bulet P.;
RT "Discovery and characterization of two isoforms of moronecidin, a
RT novel antimicrobial peptide from hybrid striped bass.";
RL J. Biol. Chem. 0:0-0(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Lauth X., Shike H., Burns J.C., Westerman M.E., Ostland V.E.,
RA Carlberg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shimizu C.,
RA Bulet P.;
RT "Discovery, and characterization of two isoforms of moronecidin, a
RT novel antimicrobial peptide from hybrid striped bass.";
RL J. Biol. Chem. 0:0-0(2002).
DR EMBL; AF332621; AAL40409.1; -
DR EMBL; AF394243; AAL57318.1; -
KW Signal.
FT CHAIN 1 22 POTENTIAL.
FT SIGNAL 23 44 MORONECIDIN.
SQ SEQUENCE 79 AA; 9106 MW; B6F32481C4200EEB CRC64;

Query Match 71.2%; Score 84; DB 13; Length 79;
Best Local Similarity 63.6%; Pred. No. 1.3e-05;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
| | | | | | | | | | : : : : :
Db 23 FFHHIFRGIVHVGKTIHKLVTG 44

RESULT 3
Q9ALA2 PRELIMINARY; PRT; 523 AA.
AC Q9ALA2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Pyridine nucleotide transhydrogenase alpha subunit.
GN PNTA.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RA Hickman J.W., Barber R.D., Skaar E., Donohue T.J.;
RT "A role for pyridine nucleotide transhydrogenase in formaldehyde
RT metabolism in Rhodobacter sphaeroides."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY026033; AAK00588.1; -
DR HSSP; Q60164; 1HZ2.
DR InterPro; IPR004002; Aladh_PNT.
DR InterPro; IPR004571; PntA.
DR Pfam; PF01262; Aladh_PNT; 1.
DR TIGRFAMS; TIGR00561; pntA; 1.
SQ SEQUENCE 523 AA; 55618 MW; 45DFCF75683EFLC9 CRC64;

Query Match 44.9%; Score 53; DB 2; Length 523;
Best Local Similarity 47.1%; Pred. No. 6.3;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 FRGIHVHAGRSIGRFLTG 22
: : : : : : : : : : : : : : :
Db 141 YRAVIEAGNNGRFFTG 157

RESULT 4
Q8UCJ1 PRELIMINARY; PRT; 524 AA.
ID Q8UCJ1
AC Q8UCJ1;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE MAD(P)+ transhydrogenase.
GN PNTA OR ATU2493 OR AGR_C_4529.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughly D., Scott C., Lapps C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009196; AAL43480.1; -
DR EMBL; AE008162; AAK88225.1; -
KW Complete proteome.
SQ SEQUENCE 524 AA; 55862 MW; 581947CD3BD68B87 CRC64;

Query Match 44.9%; Score 53; DB 16; Length 524;
Best Local Similarity 47.1%; Pred. No. 6.3;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 FRGIHVHAGRSIGRFLTG 22
: : : : : : : : : : : : : : :
Db 141 YRAVIEAGNNGRFFTG 157

RESULT 5
Q985H7 PRELIMINARY; PRT; 205 AA.
ID Q985H7
AC Q985H7;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein ml17669.
GN ML17669.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
RN EMBL; AP003012; BAB54085.1; -
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 205 AA; 22345 MW; 04F25BE0973ACC38 CRC64;

Query Match 41.5%; Score 49; DB 16; Length 205;
Best Local Similarity 44.4%; Pred. No. 9.9;
Matches 12; Conservative 1; Mismatches 8; Indels 6; Gaps 1;

OY 2 IHHIFRG-----IVHAGRSIGRFLTG 22
Db 131 IHHTAGGESAGPIQHAGEEVYLLG 157

RESULT 6

O9FR59 PRELIMINARY; PRT; 763 AA.
AC O9FR59;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Homeobox 1.
GN HB1.
OS Picea abies (Norway spruce) (Picea excelsa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX NCBI_TaxID=3329;
RN [1]
RP SEQUENCE FROM N.A.

RA Ingouff M., Farbos I., von Arnold S.;
RT "PahB1, a HD-GU2 family homeobox gene, is expressed in the outer cell
RT layer during somatic embryogenesis in Norway spruce.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF172931; AAG43405.1; -
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_repressr.
DR InterPro; IPR002913; START.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF01852; START; 1.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00234; START; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 763 AA; 83795 MW; 0EF8D5851427BAA2 CRC64;

Query Match 41.5%; Score 49; DB 10; Length 763;
Best Local Similarity 43.8%; Pred. No. 39;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 2 IHHIFRGIVHAGRSIG 17
Db 472 VHHYRQLVNSGMAFG 487

RESULT 7

O8UB08
ID O8UB08 PRELIMINARY; PRT; 83 AA.
AC O8UB08;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu3209.
GN ATU3209.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutlyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009250; AAL44025.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 9062 MW; 5392CD9518F4A0CA CRC64;

Query Match 40.7%; Score 48; DB 16; Length 83;
Best Local Similarity 52.4%; Pred. No. 5.5;
Matches 11; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

OY 2 IHHIFRGIVHAGRSIGRFLTG 22
Db 10 VHH-----GAGRRIGRFLIG 24

RESULT 8

O8ZE79 PRELIMINARY; PRT; 508 AA.
ID O8ZE79;
AC O8ZE79;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2).
GN PNTA OR YPO2303.
GN Yersinia pestis.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414151; CAC91108.1; -
DR InterPro; IPR004002; Aladh_PNT.
DR InterPro; IPR004571; PNTA.
DR Pfam; PF01262; Aladh_PNT; 1.
DR TIGRFAMS; TIGR00561; pntA; 1.
DR PROSITE; PS00836; ALADH_PNT_1; 1.

DR PROSITE; PS00837; ALADH_PNT_2; 1.
KM Oxidoreductase; Complete proteome.
SQ SEQUENCE 508 AA; 54270 MW; 8198AC8B8D82411F CRC64;

Query Match
Best Local Similarity 40.7%; Score 48; DB 16; Length 508;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGI VHAGRSIGRFLTG 22
: | | | | | | | | | |
Db 139 YRAIVEAAHEFGRRFTG 155

RESULT 9
O82PLO PRELIMINARY; PRT; 509 AA.

AC O82PLO;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pyridine nucleotide transhydrogenase (proton pump), alpha subunit (EC 1.6.1.2).
DE 1.6.1.2).
GN PNTA OR STM1479.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2."
RL Nature 413:852-856(2001).
DR EMBL; AE008764; AAL20398.1; -
DR InterPro; IPR004002; Aladh_PNT.
DR InterPro; IPR004571; PntA.
DR Pfam; PF01262; Aladh_PNT; 1.
DR TIGRFAMS; TIGR00561; pntA; 1.
DR PROSITE; PS00836; ALADH_PNT_1; 1.
DR PROSITE; PS00837; ALADH_PNT_2; 1.
KM Oxidoreductase; Complete proteome.
SQ SEQUENCE 509 AA; 54266 MW; ECC3AB4537C0D2E CRC64;

Query Match
Best Local Similarity 40.7%; Score 48; DB 16; Length 509;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGI VHAGRSIGRFLTG 22
: | | | | | | | | | |
Db 139 YRAIVEAAHEFGRRFTG 155

RESULT 10
O826X5 PRELIMINARY; PRT; 509 AA.

AC O826X5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pyridine nucleotide transhydrogenase subunit-alpha.
GN STY1589.
OS Salmonella typh.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL627270; CAD01835.1; -
DR InterPro; IPR004002; Aladh_PNT.
DR InterPro; IPR004571; PntA.
DR Pfam; PF01262; Aladh_PNT; 1.
DR TIGRFAMS; TIGR00561; pntA; 1.
DR PROSITE; PS00836; ALADH_PNT_1; 1.
DR PROSITE; PS00837; ALADH_PNT_2; 1.
KM Complete proteome.
SQ SEQUENCE 509 AA; 54236 MW; OCC39004537C0D34 CRC64;

Query Match
Best Local Similarity 40.7%; Score 48; DB 16; Length 509;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGI VHAGRSIGRFLTG 22
: | | | | | | | | | |
Db 139 YRAIVEAAHEFGRRFTG 155

RESULT 11
O9CHA3 PRELIMINARY; PRT; 883 AA.

AC O9CHA3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein yieh.
GN YIEH OR IL0835.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006317; AAK04933.1; -
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 883 AA; 99524 MW; 42ADC278C1FDC191 CRC64;

Query Match
Best Local Similarity 40.7%; Score 48; DB 16; Length 883;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGRSIGRFLT 21
| | : | | : | | : | | |
Db 51 IHSLYRNILHSGSGSGLYLT 70

RESULT 12
O9YW93 PRELIMINARY; PRT; 887 AA.

AC O9YW93;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Glycoprotein B.
GN UL27.
OS macropodid herpesvirus type 1 (parma wallaby herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
OX NCBI_TaxID=83441;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99171712; PubMed=10073704;
RA Mahony T.J., Smith G.A., Thomson D.M.;
RT "Macropodid herpesviruses 1 and 2 occupy unexpected molecular
RT phylogenetic positions within the Alphaherpesvirinae.";
RL J. Gen. Virol. 80:433-436(1999).
DR EMBL; AF061754; AAD11960.1; -
DR InterPro; IPR00234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 887 AA; 99739 MW; 0EE4FB84E098F128 CRC64;

Query Match 40.7%; Score 48; DB 12; Length 887;
Best Local Similarity 38.1%; Pred. No. 65;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGRSIGRFLTG 22
Db 718 LYSFFDGLGEIGRAVGRVVG 738

RESULT 13
Q9YW92 PRELIMINARY; PRT; 894 AA.
AC Q9YW92;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Glycoprotein B.
GN UL27.
OS macropodid herpesvirus type 2 (dorcopsis wallaby herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
OX NCBI_TaxID=83440;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99171712; PubMed=10073704;
RA Mahony T.J., Smith G.A., Thomson D.M.;
RT "Macropodid herpesviruses 1 and 2 occupy unexpected molecular
RT phylogenetic positions within the Alphaherpesvirinae.";
RL J. Gen. Virol. 80:433-436(1999).
DR EMBL; AF061755; AAD11961.1; -
DR InterPro; IPR00234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 894 AA; 100415 MW; 806C1DD45B72AECB CRC64;

Query Match 40.7%; Score 48; DB 12; Length 894;
Best Local Similarity 38.1%; Pred. No. 66;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGRSIGRFLTG 22
Db 721 LYSFFDGLGEIGRAVGRVVG 741

RESULT 14
Q8UDH4 PRELIMINARY; PRT; 357 AA.
AC Q8UDH4;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Alcohol dehydrogenase.
GN ADH OR ATU2151 OR AGR_C_3897.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009165; AAL43141.1; ALT_INIT.
DR EMBL; AE008131; AAK87898.1; -
KW Complete proteome.
SQ SEQUENCE 357 AA; 38374 MW; 68F0CE54BD358020 CRC64;

Query Match 40.3%; Score 47.5; DB 16; Length 357;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 4 HIFRGIVHA-GRSIGRFLTG 22
Db 70 HEFAGVISAAGRNVTRFKTG 89

RESULT 15
O66090 PRELIMINARY; PRT; 273 AA.
AC O66090;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Transmembrane protein Tmp5 (Fragment).
OS Lactococcus lactis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1363;
RX MEDLINE=98196737; PubMed=9537391;
RA Poquet I., Ehrlich S.D., Gruss A.;
RT "An export-specific reporter designed for gram-positive bacteria:
RT application to Lactococcus lactis.";
RL J. Bacteriol. 180:1904-1912(1998).
DR EMBL; U95840; AAC14608.1; -
FT NON_TER 273
SQ SEQUENCE 273 AA; 31411 MW; 64983CA4C554912E CRC64;

Query Match 39.8%; Score 47; DB 2; Length 273;
Best Local Similarity 45.0%; Pred. No. 27;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY	2	IHHIFRGIVHACRSIGRFLT	21
		: :	
Db	51	IHSLYRNILHSGSGQFLYT	70

Search completed: May 17, 2003, 14:27:12
Job time : 86 secs